

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
5 April 2001 (05.04.2001)

PCT

(10) International Publication Number
WO 01/23530 A1

(51) International Patent Classification⁷: C12N 5/04, 5/10,
15/10, 15/09, 15/63, 15/70, 15/74, 15/82, 15/87, C07H
21/04, A01H 1/00, 5/00, 9/00, 11/00

Alhambra Drive, #152, Davis, CA 95616 (US). DOWNIE,
Bruce; 215 Arcadia Park, Lexington, KY 40503 (US).
NONOGAKI, Hiroyuki; 408 Russel Park, #5, Davis, CA
95616 (US).

(21) International Application Number: PCT/US00/26884

(74) Agents: HYMAN, Laurence, J. et al.; Townsend and
Townsend and Crew LLP, 8th Floor, Two Embarcadero
Center, San Francisco, CA 94111-3834 (US).

(22) International Filing Date:
29 September 2000 (29.09.2000)

(25) Filing Language:

English

(81) Designated States (*national*): CA, JP.

(26) Publication Language:

English

(84) Designated States (*regional*): European patent (AT, BE,
CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC,
NL, PT, SE).

(30) Priority Data:
09/410,191 30 September 1999 (30.09.1999) US

Published:

- With international search report.
- Before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments.

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



WO 01/23530 A1

BEST AVAILABLE COPY

(54) Title: GENES EXPRESSED IN GERMINATING SEEDS AND THEIR USES

(57) Abstract: The present invention provides sequences isolated from germinating seeds that can be used to control seed germination in plants. Such enzymes include seed-specific polygalacturonases, cellulases arabinosidases, xyloglucan endotransglycosylases (XET) and expansins.

GENES EXPRESSED IN GERMINATING SEEDS AND THEIR USES

CROSS-REFERENCE TO RELATED APPLICATIONS

This application is a continuation in part of U.S. Patent Application No. 5 09/410,191, filed September 30, 1999, which is hereby incorporated by reference.

STATEMENT AS TO RIGHTS TO INVENTIONS MADE UNDER FEDERALLY SPONSORED RESEARCH AND DEVELOPMENT

This invention was made with Government support under Grant Nos. IBN-10 9407264 and UBN-9722978 awarded by the National Science Foundation. The Government has certain rights in this invention.

FIELD OF THE INVENTION

This invention relates to plant genetic engineering. In particular, it relates 15 to methods of modulating seed germination in plants

BACKGROUND OF THE INVENTION

In most seeds, radicle extension through the structures surrounding the embryo is the event that terminates germination and marks the commencement of 20 seedling growth (Bewley, *Plant Cell* 9:1055-1066 (1997)). In seeds whose embryos are embedded in a rigid endosperm, the micropylar portion of the endosperm, termed the endosperm cap, presents a physical restraint to radicle extension. This restraint must be lessened through the weakening of the endosperm cap to allow radicle emergence (Groot and Karssen, *Planta* 171: 525-531 (1987)). Endosperm cap weakening is thought to be 25 the result of cell wall hydrolysis. As mannose-containing polysaccharides are a major component of the endosperm cell walls of seeds of tomato and other Solanaceae (Sanchez *et al.*, *Plant Physiol* 93: 89-97 (1990); Dahal *et al.*, *Plant Physiol* 113: 1243-1252 (1997)), endo- β -mannanase has been regarded as a good candidate to control the 30 weakening process (Groot *et al.*, *Planta* 174: 500-504 (1988); Nomaguchi *et al.*, *Physiol Plant* 94: 105-109 (1995)). Increased mannanase activity is consistently associated with radicle emergence (Nonokaki *et al.*, *Physiol Plant* 102: 236-242 (1998)), but there are also conditions where emergence does not occur even though high enzyme activity is present (Still and Bradford, *Plant Physiol* 113: 21-29 (1997)). Thus, while endo- β -

mannanase may be necessary for germination, it does not appear to be sufficient in all cases.

In addition to endo- β -mannanase, mannosidase, galactosidase, cellulase, pectin methylesterase, polygalacturonase, arabinosidase, xyloglucan endotransglycosylase, β -1,3-glucanase and chitinase are also expressed during tomato seed germination (Groot *et al.*, *Planta* 174: 500-504 (1988); Leviatov *et al.*, *Ann Bot* 76: 1-6 (1995); Downie *et al.*, *Anal Biochem* 264: 149-157 (1998)). Since some of these hydrolases are associated with cell separation or cell wall disassembly in other developmental processes, such as abscission zones and fruit ripening (Del Campillo and Lewis, *Plant Physiol* 98: 955-961 (1992); Lashbrook *et al.*, *Plant Cell* 6: 1484-1493 (1994)), it is reasonable to expect that they also may be involved in endosperm weakening. Additional factors may also be involved in controlling this process (Bewley, *Trends Plant Sci* 2: 464-469 (1997)).

Expansins are extracellular proteins that facilitate cell wall extension, possibly by disrupting hydrogen bonding between hemicellulosic wall components and cellulose microfibrils. In addition, some expansins are expressed in non-growing tissues such as ripening fruits, where they may contribute to cell wall disassembly associated with tissue softening.

Expansin was first identified from cucumber hypocotyls by its ability to induce stress relaxation in killed cell walls (McQueen-Mason *et al.*, *Plant Cell* 4: 1425-1433 (1992). Expansins are proposed to function as cell wall loosening factors by disrupting noncovalent linkages, such as hydrogen bonds, at the cellulose-hemicellulose interface, thereby relaxing an important constraint to turgor-driven cell expansion (McQueen-Mason and Cosgrove, *Proc Natl Acad Sci USA* 91: 6574-6578 (1994); Cosgrove, *Plant Physiol* 118: 333-339 (1998)). Expansins have been highly conserved throughout plant evolution, as homologous genes have been identified in gymnosperms and in both monocots and dicots among the angiosperms. Expansins occur as multi-gene families in *Arabidopsis*, rice, cucumber, tomato, and other species where they have been examined in detail. The large number of expansin-like genes (*e.g.*, at least 22 in *Arabidopsis*) suggests multiple developmental or tissue-specific roles for these proteins, possibly in addition to vegetative growth *per se*. Expansins are expressed in shoot meristems during the early stages of leaf initiation and also in ripening fruits at a time when cell wall disassembly associated with fruit softening is occurring. During ripening,

extensive cell wall degradation and solubilization of wall components occurs, resulting in tissue softening and cell separation without cell enlargement. Preliminary results with expansin promoters linked to the GUS reporter also indicate that expression of specific expansin genes occurs in germinating seeds, in the root cap, and in association with abscission zones or tissues where cell separation will occur. Thus, in addition to their role in cell growth, specific expansins may also contribute to cell wall processes associated with developmental events such as ripening, abscission, and cell separation (Cosgrove, *Proc Natl Acad Sci USA* 94: 5504-5505 (1997)).

The prior art lacks means for controlling seed germination by controlling expression of genes associated with weakening tissues surrounding the embryo. The present invention addresses these and other needs.

SUMMARY OF THE INVENTION

The present invention provides sequences isolated from germinating seeds that can be used to control seed germination in plants. In particular, the present invention provides nucleic acid molecules which encode polypeptides having greater than about 15 70% identity to SEQ ID NO: 4, 6, 8, 10, 14, 18, 20, and 22. The polypeptides of the invention include a number of proteins and enzymes associated with weakening tissues surrounding the embryo and/or initiating radicle growth. In some embodiments, the 20 nucleic acids of the invention encode expansins that are expressed in seeds. An exemplary expansin is LeExp4 (SEQ ID NO: 1). The invention further provides nucleic acid molecules which encode polypeptides having 99% or more identity to SEQ ID NO:16, as well as nucleic acid molecules which encode polypeptides having greater than about 80% identity to SEQ ID NO:12.

The present invention further provides recombinant expression vectors comprising the nucleic acid sequences of the invention. Preferably, the vectors comprise a plant promoter operably linked to the nucleic acid sequence. The promoter can be either a constitutive promoter, or an inducible promoter.

The present invention also provides for transgenic plants comprising a 30 recombinant expression cassette of the invention. The recombinant expression cassettes are useful in methods of modulating seed germination in plants. For example, the nucleic acids of the invention can be used to enhance expression of the endogenous gene and thereby promote seed germination. Alternatively, the nucleic acids can be used to inhibit expression of the endogenous genes and thereby inhibit seed germination.

DEFINITIONS

The phrase "nucleic acid sequence" refers to a single or double-stranded polymer of deoxyribonucleotide or ribonucleotide bases read from the 5' to the 3' end. It includes chromosomal DNA, self-replicating plasmids, infectious polymers of DNA or RNA and DNA or RNA that performs a primarily structural role.

The term "promoter" refers to regions or sequence located upstream and/or downstream from the start of transcription and which are involved in recognition and binding of RNA polymerase and other proteins to initiate transcription. A "plant promoter" is a promoter capable of initiating transcription in plant cells. Such a promoter can be derived from plant genes or from other organisms, such as viruses capable of infecting plant cells.

The term "plant" includes whole plants, shoot vegetative organs/structures (e.g. leaves, stems and tubers), roots, flowers and floral organs/structures (e.g. bracts, sepals, petals, stamens, carpels, anthers and ovules), seed (including embryo, endosperm, and seed coat) and fruit (the mature ovary), plant tissue (e.g. vascular tissue, ground tissue, and the like) and cells (e.g. guard cells, egg cells, trichomes and the like), and progeny of same. The class of plants that can be used in the method of the invention is generally as broad as the class of higher and lower plants amenable to transformation techniques, including angiosperms (monocotyledonous and dicotyledonous plants), gymnosperms, ferns, and multicellular algae. It includes plants of a variety of ploidy levels, including aneuploid, polyploid, diploid, haploid and hemizygous.

A polynucleotide sequence is "heterologous to" an organism or a second polynucleotide sequence if it originates from a foreign species, or, if from the same species, is modified from its original form. For example, a promoter operably linked to a heterologous coding sequence refers to a coding sequence from a species different from that from which the promoter was derived, or, if from the same species, a coding sequence which is not naturally associated with the promoter (e.g. a genetically engineered coding sequence or an allele from a different ecotype or variety).

A polynucleotide "exogenous to" an individual plant is a polynucleotide which is introduced into the plant by any means other than by a sexual cross. Examples of means by which this can be accomplished are described below, and include *Agrobacterium*-mediated transformation, biolistic methods, electroporation, and the like. Such a plant containing the exogenous nucleic acid is referred to here as a T₁ (e.g. in

Arabidopsis by vacuum infiltration) or R₀ (for plants regenerated from transformed cells *in vitro*) generation transgenic plant. Transgenic plants that arise from sexual cross or by selfing are descendants of such a plant.

“Recombinant” refers to a human manipulated polynucleotide or a copy or complement of a human manipulated polynucleotide. For instance, a recombinant expression cassette comprising a promoter operably linked to a second polynucleotide may include a promoter that is heterologous to the second polynucleotide as the result of human manipulation (e.g., by methods described in Sambrook *et al.*, *Molecular Cloning - A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 10 (1989) or Current Protocols in Molecular Biology Volumes 1-3, John Wiley & Sons, Inc. (1994-1998)) of an isolated nucleic acid comprising the expression cassette. In another example, a recombinant expression cassette may comprise polynucleotides combined in such a way that the polynucleotides are extremely unlikely to be found in nature. For instance, human manipulated restriction sites or plasmid vector sequences may flank or 15 separate the promoter from the second polynucleotide. One of skill will recognize that polynucleotides can be manipulated in many ways and are not limited to the examples above.

The phrase "substantially identical," in the context of two nucleic acids or two polypeptides, refers to a sequence or subsequence that has at least 60% sequence identity with a reference sequence. Alternatively, percent identity can be any integer from 60% to 100%. More preferred embodiments include at least: 60%, 65%, 70%, 75%, 20 80%, 85%, 90%, 95%, or 99%. compared to a reference sequence when aligned for maximum correspondence over a comparison window as measured using one of the following sequence comparison algorithms or by manual alignment and visual inspection. Preferably, the comparison is made by BLAST using standard parameters, as described 25 below. This definition also refers to the complement of a test sequence, when the test sequence has substantial identity to a reference sequence. With respect to the LeMAN2 polypeptide, it is preferred if the sequence identity to a second polypeptide is any integer between 80% and 100%, with higher percentages of sequence identity being preferred 30 over lower percentages. With respect to the LVA-P1 polypeptide, it is preferred if the sequence identity to a second polypeptide is 99 % or 100%.

Two nucleic acid sequences or polypeptides are said to be "identical" if the sequence of nucleotides or amino acid residues, respectively, in the two sequences is the same when aligned for maximum correspondence as described below. The terms

"identical" or percent "identity," in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same, when compared and aligned for maximum correspondence over a comparison window, as measured using one of the following sequence comparison algorithms or by manual alignment and visual inspection. When percentage of sequence identity is used in reference to proteins or peptides, it is recognized that residue positions that are not identical often differ by conservative amino acid substitutions, where amino acids residues are substituted for other amino acid residues with similar chemical properties (e.g., charge or hydrophobicity) and therefore do not change the functional properties of the molecule. Where sequences differ in conservative substitutions, the percent sequence identity may be adjusted upwards to correct for the conservative nature of the substitution. Means for making this adjustment are well known to those of skill in the art. Typically this involves scoring a conservative substitution as a partial rather than a full mismatch, thereby increasing the percentage sequence identity. Thus, for example, where an identical amino acid is given a score of 1 and a non-conservative substitution is given a score of zero, a conservative substitution is given a score between zero and 1. The scoring of conservative substitutions is calculated according to, e.g., the algorithm of Meyers & Miller, *Computer Applic. Biol. Sci.* 4:11-17 (1988) e.g., as implemented in the program PC/GENE (Intelligenetics, Mountain View, California, USA).

One of skill in the art will recognize that two polypeptides can also be "substantially identical" if the two polypeptides are immunologically similar. Thus, overall protein structure may be similar while the primary structure of the two polypeptides display significant variation. Therefore a method to measure whether two polypeptides are substantially identical involves measuring the binding of monoclonal or polyclonal antibodies to each polypeptide. Two polypeptides are substantially identical if the antibodies specific for a first polypeptide bind to a second polypeptide with an affinity of at least one third of the affinity for the first polypeptide.

For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are input into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. The sequence comparison algorithm then calculates the percent sequence

identity for the test sequence(s) relative to the reference sequence, based on the designated program parameters.

Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith & Waterman, *Adv. Appl. Math.* 2:482 (1981), by the homology alignment algorithm of Needleman & Wunsch, *J. Mol. Biol.* 48:443 (1970), by the search for similarity method of Pearson & Lipman, *Proc. Nat'l. Acad. Sci. USA* 85:2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by visual inspection (*see generally, Current Protocols in Molecular Biology*, F.M. Ausubel *et al.*, eds., Current Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (1995 Supplement) (Ausubel)).

Examples of algorithms that are suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul *et al.* (1990) *J. Mol. Biol.* 215: 403-410 and Altschul *et al.* (1977) *Nucleic Acids Res.* 25: 3389-3402, respectively. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul *et al.*, *supra*). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are then extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the

sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, M=5, N=4, and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength (W) of 3, an expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff & Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915 (1989)).

5 In addition to calculating percent sequence identity, the BLAST algorithm also performs a statistical analysis of the similarity between two sequences (see, e.g., Karlin & Altschul, *Proc. Nat'l. Acad. Sci. USA* 90:5873-5787 (1993)). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)),
10 which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.1, more preferably less than about 0.01, and most preferably less than about 0.001.

15 A further indication that two nucleic acid sequences or polypeptides are substantially identical is that the polypeptide encoded by the first nucleic acid is immunologically cross reactive with the polypeptide encoded by the second nucleic acid, as described below. Thus, a polypeptide is typically substantially identical to a second polypeptide, for example, where the two peptides differ only by conservative
20 substitutions. Another indication that two nucleic acid sequences are substantially identical is that the two molecules hybridize to each other under stringent conditions, as described below.

"Conservatively modified variants" applies to both amino acid and nucleic acid sequences. With respect to particular nucleic acid sequences, conservatively modified variants refers to those nucleic acids which encode identical or essentially identical amino acid sequences, or where the nucleic acid does not encode an amino acid sequence, to essentially identical sequences. Because of the degeneracy of the genetic code, a large number of functionally identical nucleic acids encode any given protein. For instance, the codons GCA, GCC, GCG and GCU all encode the amino acid alanine.
30 Thus, at every position where an alanine is specified by a codon, the codon can be altered to any of the corresponding codons described without altering the encoded polypeptide. Such nucleic acid variations are "silent variations," which are one species of

conservatively modified variations. Every nucleic acid sequence herein which encodes a polypeptide also describes every possible silent variation of the nucleic acid. One of skill will recognize that each codon in a nucleic acid (except AUG, which is ordinarily the only codon for methionine) can be modified to yield a functionally identical molecule.

5 Accordingly, each silent variation of a nucleic acid which encodes a polypeptide is implicit in each described sequence.

As to amino acid sequences, one of skill will recognize that individual substitutions, in a nucleic acid, peptide, polypeptide, or protein sequence which alters a single amino acid or a small percentage of amino acids in the encoded sequence is a 10 "conservatively modified variant" where the alteration results in the substitution of an amino acid with a chemically similar amino acid. Conservative substitution tables providing functionally similar amino acids are well known in the art.

The following six groups each contain amino acids that are conservative substitutions for one another:

- 15 1) Alanine (A), Serine (S), Threonine (T);
 2) Aspartic acid (D), Glutamic acid (E);
 3) Asparagine (N), Glutamine (Q);
 4) Arginine (R), Lysine (K);
 5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V); and
20 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W).
*(see, e.g., Creighton, *Proteins* (1984)).*

An indication that two nucleic acid sequences or polypeptides are substantially identical is that the polypeptide encoded by the first nucleic acid is immunologically cross reactive with the antibodies raised against the polypeptide 25 encoded by the second nucleic acid. Thus, a polypeptide is typically substantially identical to a second polypeptide, for example, where the two peptides differ only by conservative substitutions. Another indication that two nucleic acid sequences are substantially identical is that the two molecules or their complements hybridize to each other under stringent conditions, as described below.

30 The phrase "selectively (or specifically) hybridizes to" refers to the binding, duplexing, or hybridizing of a molecule only to a particular nucleotide sequence under stringent hybridization conditions when that sequence is present in a complex mixture (e.g., total cellular or library DNA or RNA).

The phrase "stringent hybridization conditions" refers to conditions under which a probe will hybridize to its target subsequence, typically in a complex mixture of nucleic acid, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at 5 higher temperatures. An extensive guide to the hybridization of nucleic acids is found in Tijssen, *Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Probes*, "Overview of principles of hybridization and the strategy of nucleic acid assays" (1993). Generally, highly stringent conditions are selected to be about 5-10°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength pH. 10 Lower stringency conditions are generally selected to be about 15-30 °C below the T_m . The T_m is the temperature (under defined ionic strength, pH, and nucleic concentration) at which 50% of the probes complementary to the target hybridize to the target sequence at equilibrium (as the target sequences are present in excess, at T_m , 50% of the probes are occupied at equilibrium). Stringent conditions will be those in which the salt 15 concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (e.g., 10 to 50 nucleotides) and at least about 60°C for long probes (e.g., greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. For selective or specific 20 hybridization, a positive signal is at least two times background, preferably 10 time background hybridization.

Nucleic acids that do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides which they encode are substantially identical. This occurs, for example, when a copy of a nucleic acid is created using the 25 maximum codon degeneracy permitted by the genetic code. In such cases, the nucleic acids typically hybridize under moderately stringent hybridization conditions.

In the present invention, genomic DNA or cDNA comprising nucleic acids of the invention can be identified in standard Southern blots under stringent conditions using the nucleic acid sequences disclosed here. For the purposes of this disclosure, 30 suitable stringent conditions for such hybridizations are those which include a hybridization in a buffer of 40% formamide, 1 M NaCl, 1% SDS at 37°C, and at least one wash in 0.2X SSC at a temperature of at least about 50°C, usually about 55°C to about 60°C, for 20 minutes, or equivalent conditions. A positive hybridization is at least twice

background. Those of ordinary skill will readily recognize that alternative hybridization and wash conditions can be utilized to provide conditions of similar stringency.

Nucleic acids that do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides that they encode are substantially identical. This occurs, for example, when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code. In such cases, the nucleic acids typically hybridize under moderately stringent hybridization conditions. Exemplary "moderately stringent hybridization conditions" include a hybridization in a buffer of 40% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 1X SSC at 45°C. A positive hybridization is at least twice background. Those of ordinary skill will readily recognize that alternative hybridization and wash conditions can be utilized to provide conditions of similar stringency.

A further indication that two polynucleotides are substantially identical is if the reference sequence, amplified by a pair of oligonucleotide primers, can then be used as a probe under stringent hybridization conditions to isolate the test sequence from a cDNA or genomic library, or to identify the test sequence in, e.g., an RNA gel or DNA gel blot hybridization analysis.

DETAILED DESCRIPTION

The present invention provides nucleic acids that are expressed in seeds and can be used to control seed germination. As noted above, although endo- β -mannanase activity is consistently associated with germination, there are conditions where enzyme activity is high, yet radicle emergence does not occur. Thus, the present invention is based, at least in part on the identification of cell wall hydrolases and other enzymes that contribute to seed germination. Such enzymes include seed-specific mannanases, polygalacturonases, cellulases, arabinosidases, xyloglucan endotransglycosylases (XET) and expansins. All are dependent upon gibberellin (GA) for expression in GA-deficient seeds, are expressed initially in the endosperm caps, and are associated with cell wall hydrolysis. As *LeEXP4* is expressed specifically in endosperm caps prior to radicle emergence, this expansin is involved, possibly in conjunction with wall hydrolases, such as LeMAN2, in effecting tissue weakening.

The polypeptides encoded by the nucleic acids are thus associated with the weakening of tissues surrounding the embryo and/or initiating radicle growth. The

control of expression of the endogenous genes is therefore a convenient means for controlling seed germination. Means for controlling polypeptide activity and gene expression in plants are well known and can be used with the nucleic acids of the invention, explained below.

5 Increasing polypeptide activity or gene expression

Any of a number of means well known in the art can be used to increase activity of polypeptides or polynucleotides of the invention in plants. Enhanced expression is useful to promote seed germination. Usually isolated sequences prepared as described herein are used to prepare recombinant expression cassettes in recombinant vectors. The vectors are introduced into plant cells using methods well known to those of skill in the art. Preparation of suitable constructs and means for introducing them into plants are described below.

One of skill will recognize that the polypeptides encoded by the nucleic acids of the invention, like other proteins, have different domains that perform different functions. Thus, gene sequences of the invention need not be full length, so long as the desired functional domain of the protein is expressed.

Modified protein chains can also be readily designed utilizing various recombinant DNA techniques well known to those skilled in the art and described in detail below. For example, the chains can vary from the naturally occurring sequence at the primary structure level by amino acid substitutions, additions, deletions, and the like. These modifications can be used in a number of combinations to produce the final modified protein chain.

Alternatively, endogenous genes can be modified to enhance expression of these genes. Methods for introducing genetic mutations into plant genes and selecting plants with desired traits are well known. For instance, seeds or other plant material can be treated with a mutagenic chemical substance, according to standard techniques. Such chemical substances include, but are not limited to, the following: diethyl sulfate, ethylene imine, ethyl methanesulfonate and N-nitroso-N-ethylurea. Alternatively, ionizing radiation from sources such as, X-rays, fast neutrons or gamma rays can be used.

30 Alternatively, homologous recombination can be used to induce targeted gene modifications by specifically targeting gene of the invention *in vivo* (see, generally, Grewal and Klar, *Genetics* 146: 1221-1238 (1997) and Xu *et al.*, *Genes Dev.* 10: 2411-2422 (1996)). Homologous recombination has been demonstrated in plants (Puchta *et al.*, *Experientia* 50: 277-284 (1994), Swoboda *et al.*, *EMBO J.* 13: 484-489 (1994); Offringa

et al., Proc. Natl. Acad. Sci. USA 90: 7346-7350 (1993); and Kempin *et al. Nature* 389:802-803 (1997)).

Other means for increasing activity of polynucleotides and polypeptides of the invention can also be used. For example, one method to increase expression of genes of the invention is to use "activation mutagenesis" (see, e.g. Hiyashi *et al. Science* 258:1350-1353 (1992)). In this method an endogenous gene of the invention can be modified to be expressed constitutively, ectopically, or excessively by insertion of T-DNA sequences that contain strong/constitutive promoters upstream of the endogenous gene.

10 Inhibition of activity or expression of polynucleotides or polypeptides of the invention

Activity of endogenous genes can also be inhibited using well known techniques. Inhibition of expression of these genes can be used to inhibit seed germination and thus control the timing of seed germination, for example, by using an inducible promoter to control expression of the nucleic acids in recombinant expression 15 cassettes of the invention.

The nucleic acid sequences disclosed here can be used to design nucleic acids useful in a number of methods to inhibit expression of genes of the invention in plants. For instance, antisense technology can be conveniently used. To accomplish this, a nucleic acid segment from the desired gene is cloned and operably linked to a promoter such that the antisense strand of RNA will be transcribed. The construct is then 20 transformed into plants and the antisense strand of RNA is produced. In plant cells, it has been suggested that antisense suppression can act at all levels of gene regulation including suppression of RNA translation (see, Bourque, *Plant Sci. (Limerick)* 105: 125-149 (1995); Pantopoulos, In Progress in Nucleic Acid Research and Molecular Biology, 25 Vol. 48. Cohn, W. E. and K. Moldave (Ed.). Academic Press, Inc.: San Diego, California, USA; London, England, UK. p. 181-238; Heiser *et al., Plant Sci. (Shannon)* 127: 61-69 (1997)) and by preventing the accumulation of mRNA which encodes the protein of interest, (see, Baulcombe, *Plant Mol. Bio.* 32:79-88 (1996); Prins and Goldbach, *Arch. Virol.* 141: 2259-2276 (1996); Metzlaff *et al. Cell* 88: 845-854 (1997), Sheehy *et al.*, *Proc. Natl. Acad. Sci. USA*, 85:8805-8809 (1988), and Hiatt *et al.*, U.S. Patent No. 30 4,801,340).

The nucleic acid segment to be introduced generally will be substantially identical to at least a portion of the endogenous gene or genes to be repressed. The sequence, however, need not be perfectly identical to inhibit expression. The vectors of

the present invention can be designed such that the inhibitory effect applies to other genes within a family of genes exhibiting identity or substantial identity to the target gene.

For antisense suppression, the introduced sequence also need not be full length relative to either the primary transcription product or fully processed mRNA.

- 5 Generally, higher identity can be used to compensate for the use of a shorter sequence. Furthermore, the introduced sequence need not have the same intron or exon pattern, and identity of non-coding segments may be equally effective. Normally, a sequence of between about 30 or 40 nucleotides to about the full length of a nucleotide should be used, though a sequence of at least about 100 nucleotides is preferred, a sequence of at 10 least about 200 nucleotides is more preferred, and a sequence of about 500 to about 3500 nucleotides is especially preferred.

A number of gene regions can be targeted to suppress expression of genes of the invention. The targets can include, for instance, the coding regions, introns, sequences from exon/intron junctions, 5' or 3' untranslated regions, and the like.

- 15 Another well-known method of suppression is sense co-suppression. Introduction of nucleic acid configured in the sense orientation has been recently shown to be an effective means by which to block the transcription of target genes. For an example of the use of this method to modulate expression of endogenous genes (see, Assaad *et al.*, *Plant Mol. Bio.* 22: 1067-1085 (1993); Flavell, *Proc. Natl. Acad. Sci. USA* 20 91: 3490-3496 (1994); Stam *et al.*, *Annals Bot.* 79: 3-12 (1997); Napoli *et al.*, *The Plant Cell* 2:279-289 (1990); and U.S. Patents Nos. 5,034,323, 5,231,020, and 5,283,184).

- 25 The suppressive effect may occur where the introduced sequence contains no coding sequence *per se*, but only intron or untranslated sequences homologous to sequences present in the primary transcript of the endogenous sequence. The introduced sequence generally will be substantially identical to the endogenous sequence intended to be repressed. This minimal identity will typically be greater than about 65%, but a higher identity might exert a more effective repression of expression of the endogenous sequences. Substantially greater identity of more than about 80% is preferred, though about 95% to absolute identity is most preferred. As with antisense regulation, the effect 30 should apply to any other proteins within a similar family of genes exhibiting identity or substantial identity.

For co-suppression, the introduced sequence, needing less than absolute identity, also need not be full length, relative to either the primary transcription product or fully processed mRNA. This may be preferred to avoid concurrent production of some

plants that over-express the introduced sequence. A higher identity in a sequence shorter than full-length compensates for a longer, less identical sequence. Furthermore, the introduced sequence need not have the same intron or exon pattern, and identity of non-coding segments will be equally effective. Normally, a sequence of the size ranges noted 5 above for antisense regulation is used. In addition, the same gene regions noted for antisense regulation can be targeted using co-suppression technologies.

Other means of inhibiting expression are known. These methods include formation of triple-helix DNA (see, e.g., Havre and Glazer *J. Virology* 67:7324-7331 10 (1993); Scanlon *et al.* *FASEB J.* 9:1288-1296 (1995); Giovannangeli *et al.* *Biochemistry* 35:10539-10548 (1996); Chan and Glazer *J. Mol. Medicine (Berlin)* 75: 267-282 (1997)) and ribozymes (Zhao and Pick, *Nature* 365:448-451 (1993); Eastham and Ahlering, *J. Urology* 156:1186-1188 (1996); Sokol and Murray, *Transgenic Res.* 5:363-371 (1996); Sun *et al.*, *Mol. Biotechnology* 7:241-251 (1997); and Haseloff *et al.*, *Nature*, 334:585-591 (1988)).

15 Modification of endogenous genes can also be used to inhibit expression. Methods for introducing genetic mutations described above can also be used to select for plants with decreased expression of genes of the invention.

20 Other means for inhibiting polynucleotide or polypeptide activity can also be used. Activity of polynucleotides of the invention may be modulated by eliminating the proteins that are required for cell-specific expression of such polynucleotides. Thus, expression of regulatory proteins and/or the sequences that control gene expression can be modulated using the methods described here.

Purification of polypeptides

25 Naturally occurring or recombinant polypeptides of the invention can be purified for use in functional assays. Naturally occurring polypeptides can be purified, e.g., from plant tissue and any other source of the desired polypeptide. Recombinant polypeptides can be purified from any suitable expression system.

30 The polypeptides of the invention may be purified to substantial purity by standard techniques, including selective precipitation with such substances as ammonium sulfate; column chromatography, immunopurification methods, and others (see, e.g., Scopes, *Protein Purification: Principles and Practice* (1982); U.S. Patent No. 4,673,641; Ausubel *et al.*, *supra*; and Sambrook *et al.*, *supra*).

A number of procedures can be employed when recombinant polypeptides are being purified. For example, proteins having established molecular adhesion

properties (e.g. epitope tags, histidine tags and the like) can be reversibly fused to polypeptides of the invention. With the appropriate ligand, the such polypeptides can be selectively adsorbed to a purification column and then freed from the column in a relatively pure form.

5 Isolation of nucleic acids of the invention

Generally, the nomenclature and the laboratory procedures in recombinant DNA technology described below are those well known and commonly employed in the art. Standard techniques are used for cloning, DNA and RNA isolation, amplification and purification. Generally enzymatic reactions involving DNA ligase, DNA polymerase, 10 restriction endonucleases and the like are performed according to the manufacturer's specifications. These techniques and various other techniques are generally performed according to Sambrook *et al.*, *Molecular Cloning - A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, (1989) or Current Protocols in Molecular Biology Volumes 1-3, John Wiley & Sons, Inc. (1994-1998).

15 The isolation of nucleic acids of the invention may be accomplished by a number of techniques. For instance, oligonucleotide probes based on the sequences disclosed here can be used to identify the desired gene in a cDNA or genomic DNA library. To construct genomic libraries, large segments of genomic DNA are generated by random fragmentation, e.g. using restriction endonucleases, and are ligated with vector 20 DNA to form concatemers that can be packaged into the appropriate vector. To prepare a cDNA library, mRNA is isolated from the desired organ, such as seeds, and a cDNA library that contains a gene transcript of the invention is prepared from the mRNA. Alternatively, cDNA may be prepared from mRNA extracted from other tissues in which genes of the invention or homologs are expressed.

25 The cDNA or genomic library can then be screened using a probe based upon the sequence of a cloned gene of the invention as disclosed here. Probes may be used to hybridize with genomic DNA or cDNA sequences to isolate homologous genes in the same or different plant species. Alternatively, antibodies raised against a polypeptide of the invention can be used to screen an mRNA expression library.

30 Alternatively, the nucleic acids of interest can be amplified from nucleic acid samples using amplification techniques. For instance, polymerase chain reaction (PCR) technology can be used to amplify the sequences of genes of the invention directly from genomic DNA, from cDNA, from genomic libraries or cDNA libraries. PCR and other *in vitro* amplification methods may also be useful, for example, to clone nucleic

acid sequences that code for proteins to be expressed, to make nucleic acids to use as probes for detecting the presence of the desired mRNA in samples, for nucleic acid sequencing, or for other purposes. For a general overview of PCR, see *PCR Protocols: A Guide to Methods and Applications*. (Innis, M, Gelfand, D., Sninsky, J. and White, T., eds.), Academic Press, San Diego (1990). Appropriate primers and probes for identifying sequences of the invention from plant tissues are generated from comparisons of the sequences provided here (e.g. SEQ ID NO: 1, SEQ ID NO:3, etc.).

5 Polynucleotides may also be synthesized by well-known techniques as described in the technical literature. See, e.g., Carruthers *et al.*, *Cold Spring Harbor Symp. Quant. Biol.* 47:411-418 (1982), and Adams *et al.*, *J. Am. Chem. Soc.* 105:661 (1983). Double stranded DNA fragments may then be obtained either by synthesizing the complementary strand and annealing the strands together under appropriate conditions, or by adding the complementary strand using DNA polymerase with an appropriate primer sequence.

15

Preparation of recombinant vectors

To use isolated sequences in the above techniques, recombinant DNA vectors suitable for transformation of plant cells are prepared. Techniques for transforming a wide variety of higher plant species are well known and described in the technical and scientific literature. See, for example, Weising *et al.* *Ann. Rev. Genet.* 22:421-477 (1988). A DNA sequence coding for the desired polypeptide, for example a cDNA sequence encoding a full length protein, will preferably be combined with transcriptional and translational initiation regulatory sequences which will direct the transcription of the sequence from the gene in the intended tissues of the transformed plant.

20 For example, for overexpression, a plant promoter fragment may be employed which will direct expression of the gene in all tissues of a regenerated plant. Such promoters are referred to herein as "constitutive" promoters and are active under most environmental conditions and states of development or cell differentiation. Examples of constitutive promoters include the cauliflower mosaic virus (CaMV) 35S transcription initiation region, the 1'- or 2'- promoter derived from T-DNA of *Agrobacterium tumefaciens*, and other transcription initiation regions from various plant genes known to those of skill. Constitutive promoters and regulatory elements can also be isolated from genes that are expressed constitutively or at least expressed in most if not

all tissues of a plant. Such genes include, for example, *ACT11* from *Arabidopsis* (Huang et al. *Plant Mol. Biol.* 33:125-139 (1996)), *Cat3* from *Arabidopsis* (GenBank No. U43147, Zhong et al., *Mol. Gen. Genet.* 251:196-203 (1996)), the gene encoding stearoyl-acyl carrier protein desaturase from *Brassica napus* (Genbank No. X74782, Solocombe et al. *Plant Physiol.* 104:1167-1176 (1994)), *Gpc1* from maize (GenBank No. X15596, Martinez et al. *J. Mol. Biol.* 208:551-565 (1989)), and *Gpc2* from maize (GenBank No. U45855, Manjunath et al., *Plant Mol. Biol.* 33:97-112 (1997)).

Alternatively, the plant promoter may direct expression of a nucleic acid of the invention in a specific tissue, organ or cell type (*i.e.* tissue-specific promoters) or may be otherwise under more precise environmental or developmental control (*i.e.* inducible promoters). Examples of environmental conditions that may effect transcription by inducible promoters include anaerobic conditions, elevated temperature, the presence of light, or application of chemicals/hormones (such promoters can be used, for example, in the chemical induction of antisense sequences for inhibition of seed germination until a desired time). Exemplary promoters for this purpose include promoters from glucocorticoid receptor genes (Aoyama and Chau, *Plant J* 11:605-12 (1997)). Tissue-specific promoters may only promote transcription within a certain time frame of developmental stage within that tissue. Other tissue specific promoters may be active throughout the life cycle of a particular tissue. One of skill will recognize that a tissue-specific promoter may drive expression of operably linked sequences in tissues other than the target tissue. Thus, as used herein a tissue-specific promoter is one that drives expression preferentially in the target tissue or cell type, but may also lead to some expression in other tissues as well.

A number of tissue-specific promoters can be used in the invention. Preferred promoters are those that direct expression of nucleic acids in seeds. As used herein a seed-specific promoter is one which directs expression in seed tissues, such promoters may be, for example, ovule-specific, embryo-specific, endosperm-specific, integument-specific, seed coat-specific, or some combination thereof. Examples include a promoter from the ovule-specific *BEL1* gene described in Reiser et al. *Cell* 83:735-742 (1995) (GenBank No. U39944). Other suitable seed specific promoters are derived from the following genes: *MAC1* from maize (Sheridan et al. *Genetics* 142:1009-1020 (1996), *Cat3* from maize (GenBank No. L05934, Abler et al. *Plant Mol. Biol.* 22:10131-1038 (1993), the gene encoding oleosin 18kD from maize (GenBank No. J05212, Lee et al. *Plant Mol. Biol.* 26:1981-1987 (1994)), viviparous-1 from *Arabidopsis* (Genbank No.

U93215), the gene encoding oleosin from *Arabidopsis* (Genbank No. Z17657), Atmyc1 from *Arabidopsis* (Urao *et al.* *Plant Mol. Biol.* 32:571-576 (1996)), the 2s seed storage protein gene family from *Arabidopsis* (Conceicao *et al.* *Plant* 5:493-505 (1994)) the gene encoding oleosin 20kD from *Brassica napus* (GenBank No. M63985), *napA* from 5 *Brassica napus* (GenBank No. J02798, Josefsson *et al.* *JBL* 26:12196-1301 (1987)), the *napin* gene family from *Brassica napus* (Sjodahl *et al.* *Planta* 197:264-271 (1995)), the gene encoding the 2S storage protein from *Brassica napus* (Dasgupta *et al.* *Gene* 133:301-302 (1993)), the genes encoding oleosin A (Genbank No. U09118) and oleosin B (Genbank No. U09119) from soybean and the gene encoding low molecular weight 10 sulphur rich protein from soybean (Choi *et al.* *Mol Gen. Genet.* 246:266-268 (1995)).

Another example of a promoter useful in the present invention is the promoter of the *LeXPG1* gene provided in SEQ ID NO: 23. One of skill will recognize that the variants of this promoter sequence can also be used. For example, the promoter can be less than full length (e.g. fragments of 500 to about 1000 nucleotides in length) 15 and still provide suitable expression levels.

The vector comprising the sequences (e.g., promoters or coding regions) from genes of the invention will typically comprise a marker gene that confers a selectable phenotype on plant cells. For example, the marker may encode biocide resistance, particularly antibiotic resistance, such as resistance to kanamycin, G418, 20 bleomycin, hygromycin, or herbicide resistance, such as resistance to chlorosulfuron or Basta.

Production of transgenic plants

DNA constructs of the invention may be introduced into the genome of the desired plant host by a variety of conventional techniques. For example, the DNA 25 construct may be introduced directly into the genomic DNA of the plant cell using techniques such as electroporation and microinjection of plant cell protoplasts, or the DNA constructs can be introduced directly to plant tissue using ballistic methods, such as DNA particle bombardment.

Microinjection techniques are known in the art and well described in the 30 scientific and patent literature. The introduction of DNA constructs using polyethylene glycol precipitation is described in Paszkowski *et al.* *EMBO. J.* 3:2717-2722 (1984). Electroporation techniques are described in Fromm *et al.* *Proc. Natl. Acad. Sci. USA* 82:5824 (1985). Ballistic transformation techniques are described in Klein *et al.* *Nature* 327:70-73 (1987).

Alternatively, the DNA constructs may be combined with suitable T-DNA flanking regions and introduced into a conventional *Agrobacterium tumefaciens* host vector. The virulence functions of the *Agrobacterium tumefaciens* host will direct the insertion of the construct and adjacent marker into the plant cell DNA when the cell is infected by the bacteria. *Agrobacterium tumefaciens*-mediated transformation techniques, including disarming and use of binary vectors, are well described in the scientific literature. See, for example Horsch *et al.* *Science* 233:496-498 (1984), and Fraley *et al.* *Proc. Natl. Acad. Sci. USA* 80:4803 (1983) and *Gene Transfer to Plants*, Potrykus, ed. (Springer-Verlag, Berlin 1995).

Transformed plant cells which are derived by any of the above transformation techniques can be cultured to regenerate a whole plant which possesses the transformed genotype and thus the desired phenotype such as increased seed mass. Such regeneration techniques rely on manipulation of certain phytohormones in a tissue culture growth medium, typically relying on a biocide and/or herbicide marker that has been introduced together with the desired nucleotide sequences. Plant regeneration from cultured protoplasts is described in Evans *et al.*, *Protoplasts Isolation and Culture, Handbook of Plant Cell Culture*, pp. 124-176, MacMillan Publishing Company, New York, 1983; and Binding, *Regeneration of Plants, Plant Protoplasts*, pp. 21-73, CRC Press, Boca Raton, 1985. Regeneration can also be obtained from plant callus, explants, organs, or parts thereof. Such regeneration techniques are described generally in Klee *et al.* *Ann. Rev. of Plant Phys.* 38:467-486 (1987).

The nucleic acids of the invention can be used to confer desired traits on essentially any plant. Thus, the invention has use over a broad range of plants, including species from the genera *Anacardium*, *Arachis*, *Asparagus*, *Atropa*, *Avena*, *Brassica*, *Citrus*, *Citrullus*, *Capsicum*, *Carthamus*, *Cocos*, *Coffea*, *Cucumis*, *Cucurbita*, *Daucus*, *Elaeis*, *Fragaria*, *Glycine*, *Gossypium*, *Helianthus*, *Heterocallis*, *Hordeum*, *Hyoscyamus*, *Lactuca*, *Linum*, *Lolium*, *Lupinus*, *Lycopersicon*, *Malus*, *Manihot*, *Majorana*, *Medicago*, *Nicotiana*, *Olea*, *Oryza*, *Panicum*, *Pannisetum*, *Persea*, *Phaseolus*, *Pistacia*, *Pisum*, *Pyrus*, *Prunus*, *Raphanus*, *Ricinus*, *Secale*, *Senecio*, *Sinapis*, *Solanum*, *Sorghum*, *Theobromus*, *Trigonella*, *Triticum*, *Vicia*, *Vitis*, *Vigna*, and *Zea*.

One of skill will recognize that after the expression cassette is stably incorporated in transgenic plants and confirmed to be operable, it can be introduced into other plants by sexual crossing. Any of a number of standard breeding techniques can be used, depending upon the species to be crossed.

Using known procedures one of skill can screen for plants of the invention by detecting the increase or decrease of mRNA or protein of the invention in transgenic plants, particularly in the seed. Means for detecting and quantitating mRNAs or proteins are well known in the art.

5 Plants with modulated seed germination can be easily be selected by monitoring seed germination using standard techniques.

EXAMPLES

The following examples are offered to illustrate, but not to limit the claimed invention.

10

Example 1:

This Example describes the cloning of expansin genes from tomato seeds.

MATERIALS AND METHODS

15

Plant Materials

Tomato (*Lycopersicon esculentum* Mill.) seeds from either wild-type (cv. Moneymaker) plants or homozygous gibberellin-deficient (*gib-1*) mutant plants were used throughout the study. The *gib-1* mutant and its isogenic parent line were originally obtained from Dr. Cees Karssen, Wageningen Agricultural University, The Netherlands. Mutant plants were sprayed 3 times per week with 100 µM GA to revert the dwarf habit and allow more vigorous growth and fertility. After fruits were harvested, seeds were extracted, treated with 0.25 M HCl, dried to 6% moisture content (fresh basis) and stored at -20°C until used (Ni and Bradford, *Plant Physiol* 101: 607-617 (1993)).

25

Germination Conditions

Approximately 500 seeds were incubated at 25°C in the dark in 9-cm diameter Petri dishes on top of two layers of filter paper moistened with 12 mL of deionized water, 100 µM GA₄₊₇, 100 µM ABA or PEG 8000 solutions having water potentials of -0.5, -1.0 and -2.0 MPa. For far-red (FR) light treatment, seeds were imbibed at 25°C for 40 h under continuous FR illumination in a custom-made FR chamber (Lagarias *et al.*, *Plant Cell* 9: 675-688 (1997) where peak transmittance, half-band pass, and fluence rate were 760 nm, 85 nm and 22 µmol m⁻²s⁻¹, respectively, at the

level of the seed as measured by a LI-COR LI-8000 portable spectroradiometer (LI-COR, Inc., Lincoln, Nebraska, USA).

RNA Isolation, PCR Amplification, and cDNA Library Screening

5 Samples of 500 whole seeds imbibed for 24 h were pulverized in LN₂ and the frozen material transferred to 2 mL of extraction buffer (10 mM Tris-HCl pH 8.2, 100 mM LiCl, 1 mM EDTA, 1% (w/v) SDS, 25 mM DTT) in a ground glass homogenizer on ice. Extraction followed a modification of the phenol/SDS method of Ausubel *et al.* Current Protocols in Molecular Biology. (New York: Wiley-Interscience, 1987). One µg 10 of purified total RNA was used as template for RT-PCR. Two degenerate PCR primers, the 5' primer [G(GC)(N)CA(TC)GC(N)AC(N)TT(CT)TA(CT)GG(N)G] and the 3' primer [(TC)TGCCA(AG)TT(TC)TG(N)CCCCA(AG)TT] (N=A, T, G or C) were designed based on two conserved amino acid domains according to the alignment of deduced 15 amino acid sequences of known expansins (Shcherban *et al.*, *Proc Natl Acad Sci USA* 92: 9245-9249 (1995); Cho and Kende *Plant Cell* 9:1661-1671 (1997); Rose *et al.*, *Proc Natl Acad Sci USA* 94: 5955-5960 (1997). After amplification for 36 cycles (94°C for 1 min, 50°C for 1.5 min, and 72°C for 1.5 min), the amplified fragments were cloned into pCR2.1 according to the manufacturer's instructions (Invitrogen). DNA sequences were determined with universal primers (T3 and M13-forward) using an Applied Biosystems 20 model 377 sequencer (Perkin-Elmer) with dye termination chemistry and AmpliTaq DNA polymerase FS (Perkin-Elmer/Applied Biosystems). The PCR fragments were used to screen a cDNA library prepared from *gib-1* seeds imbibed in 100 µM GA₄₊₇ for 24 h. The cDNAs were labeled with enhanced chemiluminescence (ECL) labeling reagents (Amersham Life Science, Inc.), then were added to prehybridization solution at a final 25 concentration of 10 ng/mL. Prehybridization was for 30 min at 42°C, and hybridization was for 3 h at 42°C. Following hybridization, the membranes were washed twice at 42°C with 6 M urea, 0.5% SDS (low stringency) or 0.2% SDS (high stringency) and then washed twice for 5 min each with 2X SSC at room temperature. Independent inserts in the library vector pBK-CMV were sequenced.

30

DNA Gel Blot Analyses

For cDNA gel blot analysis, 5 ng of PCR product obtained by T3/T7 primers from library vector containing target gene was subjected to electrophoresis, and

transferred to a Hybond-N⁺ membrane (Amersham Life Science, Inc.). Prehybridization, hybridization, washing, and autoradiography of the blot were performed as described for cDNA library screening using ECL lit, except that hybridization was kept for overnight.

Genomic DNA was isolated from young tomato leaves (cv. Moneymaker) as described by Murray and Thompson *Nucleic Acids Res* 8: 4321-4325 (1980) and modified by Bernatzky and Tanksley *Theor Appl Genet* 72: 314-321 (1986). Aliquots (10 µg) were digested with restriction enzymes, fractionated on a 0.8% agarose gel, and transferred to Hybond-N⁺ membranes. Prehybridization, hybridization, washing, and autoradiography of the blots were performed as described for cDNA gel blot analysis.

10

RNA Gel Blot Analyses

Total RNA was isolated from germinating seeds, seed parts or different tissues according to the above method. For seed part RNA extraction, imbibed seeds were first dissected into three parts: the endosperm cap, the radicle tip removed from within the endosperm cap, and the rest of seed (lateral endosperm and remainder of embryo). Three pools containing 1000 seed parts were used for RNA isolation. Total RNA from each sample (5 µg) was subjected to electrophoresis on 1% (w/v) agarose/10% formaldehyde denaturing gels, transferred to Hybond-N⁺ membrane and UV cross-linked. The probe for Northern Blots was DIG-labeled DNA probe prepared by using PCR labeling method. Primers corresponding to 3'-terminal untranslation region of the gene were used to incorporate DIG-labeled dNTP into DNA probe. The labeling efficiency was estimated according to the manual (Boehringer Mannheim, Inc.). Final probe concentration used for hybridization is 25ng/ML hybridization buffer. Since DNA probe is used for RNA detection, high SDS buffer (7% SDS) was used for hybridization at 42⁰C. The washing (60⁰C) and detection followed the recommended method using the chemiluminescent substrate CPSD (Boehringer Mannheim, Inc.). Exposure time was from 10 min to 2 h depending on the strength of the signal.

Puncture Force Measurements

30 The force required to puncture the micropylar endosperm and testa surrounding the radicle tip was analyzed for both *gib-1* and MM seeds following various treatments. Each micropylar tip was sliced from the seed and the radicle tip teased out of the embryo cavity. A food texture analyzer (Stable Micro Systems, Texture Technologies Corp., Scarsdale, New York, USA) fitted with a custom-made probe (0.5 mm diameter)

was used to determine puncture force. The endosperm cap without the radicle tip was placed on the texture analyzer probe and the test conducted at an inching speed of 10 mm/min (Downie, et al., 1999). In each test, the background resistance generated by the probe against the side of the cap was subtracted from the peak resistance to puncture force using XT.RA Dimension version 3.7F software supplied by the manufacturer (Stable Micro Systems). Twenty-four individual seeds were measured at each time point and means were expressed as the puncture force in newtons (N).

RESULTS

10 *Cloning and Sequence Analysis of Tomato Seed Expansins*

Using primers to conserved expansin sequences, a ~540 bp cDNA band was amplified by RT-PCR from germinating tomato seed RNA and subsequent sequence analyses indicated the existence of six expansin homologs (termed TE1 to TE6). These 15 six independent fragments share high amino acid sequence identity among themselves and with two truncated cucumber expansins from hypocotyl. The predicted amino acid sequence of all six tomato expansin homologs exhibit the basic features of expansins, including eight conserved cysteines and conserved tryptophans (Shcherban *et al.*, 1995). The sequence of TE1 was identical to truncated sequence of *LeExp1*, which is expressed 20 during tomato fruit ripening (Rose *et al.*, 1997), and the sequence of TE2 was identical to truncated sequence of *LeExp2*, which is expressed in expanding tissues (Reinhardt *et al.*, *Plant Cell* 10: 1427-1438 (1998)). TE3 was identical with the truncated sequence of *LeExp4* (SEQ ID NO: 1) which is detected in flowers (Brummell *et al.*, *Plant Mol Biol* 39: 161-169 (1999)). The remaining three fragments TE4, TE5 and TE6 were unique in 25 the database. Since the colonies were randomly selected for sequencing, it is possible that additional expansin homologs are expressed in germinating tomato seeds.

A cDNA library prepared from GA-treated *gib-1* tomato seeds was screened with TE1 to TE6 respectively. The full length sequences for TE1, TE2 and TE3 were confirmed to be identical with *LeExp1*, *LeExp2* and *LeExp4*. The full length 30 sequences for TE4, TE5 and TE6 were named *LeExp8*, (SEQ ID NO: 3) *LeExp9* and *LeExp10* (SEQ ID NO: 5), respectively. Database searching with the deduced amino acid sequences of the full-length cDNAs revealed high degree of homology to expansins from cucumber, *Arabidopsis*, rice, pea and tomato.

LeExp4 is a Member of a Multigene Family with Tissue-Specific Expression in Germinating Seeds

5 To test expression level of individual expansin genes, a gene-specific probe was designed based on 3'- terminal untranslation region and obtained by PCR and then used for Northern Blot analysis. No signal or very weak signal can be detected from germinating seeds for *LeExp1*, *LeExp2* or *LeExp9*, while strong signal can be detected for *LeExp4*, *LeExp8* and *LeExp10*.

10 *Hormonal and Environmental Regulation of LeExp4 Expression and Endosperm Cap Weakening*

Endosperm cap weakening and germination of gibberellin-deficient *gib-1* tomato seeds is dependent on GA. To determine the expression pattern of *LeExp4* in response to GA, total RNA was extracted from *gib-1* seeds after imbibition for 2, 12, 24, 15 48 and 60h in water or in 100 µM GA and hybridized with the *LeExp4*-specific RNA probe. No expression was detected in dry *gib-1* seeds or *gib-1* seeds imbibed in water. In the presence of GA, *LeExp4* transcript accumulated within 12 h and reached a maximal level at 24 h of imbibition. Subsequently, *LeExp4* mRNA abundance declined slightly by 20 48 h. (Approximately 4% of the seeds had completed radicle emergence by 48 h, although only ungerminated seeds were sampled for RNA). Puncture force analysis showed that endosperm caps did not weaken when *gib-1* seeds were imbibed in water, but weakening did occur when GA was present in the imbibition solution. Weakening was evident 25 within 12 h of imbibition in GA, a time when *LeExp4* mRNA was also detected.

Germination of wild-type tomato seeds can be inhibited by ABA. The effect of ABA is similar in MM seeds with *gib-1* seeds. When total RNA was extracted from MM seeds after imbibition for 2, 12, 24, 48 and 60 h in water or in 100 µM ABA and hybridized with the *LeExp4*-specific probe. *LeExp4* mRNA could be detected from 30 12 h of imbibition in both water and ABA, and the pattern of expression was similar regardless of the presence of ABA. ABA also had no effect on the decrease in puncture force of the endosperm cap, although it did prevent radicle emergence.

Low water potential can delay or prevent tomato seed germination. When tomato seeds were imbibed in PEG solutions maintaining -0.5 or -1.0 MPa osmotic

potentials, the abundance of *LeExp4* mRNA slightly decreased as the water potential decreased. Puncture force analysis also showed that low water potential can prevent or delay the endosperm weakening process.

Far-red light can also inhibit tomato seed germination. When MM seeds were imbibed under continuous far-red light, germination was prevented and no expression of *LeExp4* was detected. Far-red light also prevented any weakening of the endosperm cap.

DISCUSSION

Expansins comprise a large superfamily of genes sharing conserved sequences both within and among species. The results presented here show here that *LeExp4* has a highly specific pattern of expression. Using a gene-specific probe, *LeExp4* mRNA was detected in flowers and germinating seeds. In the germinating seed, expression of *LeExp4* was restricted to the endosperm cap tissue directly opposite the radicle tip. This endosperm cap tissue is anatomically and physiologically differentiated from the remainder of the endosperm, for example by having thinner cell walls and by the expression of a cap-specific isoform of endo- β -mannanase prior to radicle emergence. Cell wall degradation, vacuolization, and other visible changes occur in the cap region prior to the initiation of such changes in the remainder of the endosperm. Thus, the occurrence of a specific expansin and of unique isoforms of hydrolytic enzymes in the endosperm cap is likely associated with the tissue weakening and cell separation that occurs to allow radicle emergence, rather than being strictly a reserve mobilization phenomenon as in the remainder of the endosperm.

This conclusion is supported by the expression pattern of *LeExp4* in response to factors that also regulate germination. In *gib-1* seeds, which require GA to complete germination, GA induces accumulation of *LeExp4* mRNA within 12 h of imbibition. Even though the entire seed is in contact with the imbibition solution containing GA, *LeExp4* mRNA is detected only in the endosperm cap. On the other hand, both reduced water potential and far-red light inhibit germination reduce or prevent expression of *LeExp4*. Furthermore, the extent of endosperm cap weakening was quantitatively related to the abundance of *LeExp4* mRNA.

The relationship of *LeExp4* expression to ABA at first appears rather anomalous. ABA effectively inhibits germination, and in isolated endosperm caps it also

blocks GA-induced weakening. ABA did not prevent *LeExp4* expression and had no effect on endosperm weakening. Thus, ABA does not appear to inhibit germination by blocking the expression of genes associated with endosperm cap weakening. It has been suggested that there is a second phase of weakening required for radicle emergence, and 5 that ABA inhibits this second phase. Alternatively, ABA may act primarily on the growth potential of the embryo, reducing it below that required to penetrate even the weakened endosperm cap.

The expansin gene *LeExp4* is expressed specifically in the endosperm cap of imbibed tomato seeds. Expression of *LeExp4* is initiated within 12 h of imbibition, is 10 regulated by factors that affect germination, and is quantitatively correlated with the extent of weakening of the endosperm cap tissues. Together, these results support the hypothesis that *LeExp4*, most likely in conjunction with cell wall hydrolases, is involved in the cell wall changes associated with tissue weakening and cell separation in the endosperm cap. If expansin protein is required to loosen hemicellulosic bonds and/or to 15 allow access of hydrolases to the polymer matrix, regulation of its expression could be a critical control point in the germination process.

Example 2

Using the RT-PCR approach described above, other cDNAs from genes 20 associated with germination were isolated. For example polygalacturonase (PG) that is expressed in imbibed tomato seeds was identified. Extracts from seeds have exhibited only exo-PG activity and no endo-PG activity, this enzyme is likely encoded by the cloned gene (*LeXPG 1*, SEQ ID NO: 21; the encoded polypeptide is set forth in SEQ ID NO:22). cDNAs (*Cel55* and *Cel68* SEQ ID NO: 17 and 19) having high homology to 25 known β-1,4-glucanases have also been identified. The encoded polypeptide sequences are set forth in SEQ ID NOs:18 and 20, respectively.

A third putative hydrolase cDNA (*LeARA 1*, SEQ ID NO: 9, the encoded 30 polypeptide is set forth in SEQ ID NO:10) was isolated from a differential cDNA display screen of *gib-1* mutant tomato seeds imbibed in the presence or absence of GA. The corresponding mRNA was initially expressed only in the endosperm caps in response to GA, and subsequently expression spread through the remainder of the endosperm, but not into the embryo. The predicted amino acid sequence showed high homology to bacterial and fungal arabinosidases, and seed extracts exhibited arabinosidase activity.

Another candidate hydrolase involved in germination is xyloglucan endotransglycosylase (XET, SEQ ID NO: 13, the encoded polypeptide is set forth as SEQ ID NO:14). It is believed that xyloglucans, the principal hemicellulose component in the primary cell walls of dicots, can form a tightly bound, non-covalent association with cellulose. XET catalyzes both the endo-type splitting of a xyloglucan molecule and the linking of the newly generated reducing end to a nonreducing end of another xyloglucan molecule or an oligosaccharide. This lengthening and rearrangement of xyloglucans may release tension and accommodate wall expansion during cell growth. Other genes identified in this way include a gibberellin-stimulated/ABA down regulated gene of unknown function (LeGAS2, SEQ ID NO: 7, the encoded polypeptide is set forth as SEQ ID NO:8).

Example 3

This example described the cloning of a mannanase gene (LeMAN2, SEQ ID NO: 11) from tomato seeds.

MATERIALS AND METHODS

Plant Material and Seed Germination

Tomato (*Lycopersicon esculentum* Mill.) seeds, either from wild-type (cv Moneymaker) plants or homozygous gibberellin-deficient mutant plants (*gib-1*) were used in this study. Mutant plants were sprayed three times per week with 100 µM GA₄₊₇ to allow more vigorous growth and fertility. After fruits were harvested, seeds were collected, treated with 0.25 M HCl, dried to 6% moisture content (fresh weight basis) and stored at 20°C until used (Ni, B.R. *et al.*, *Plant Physiol* 101:607-617 (1993)). For germination, 100 or 200 tomato seeds were placed on two layers of filter papers moistened with 12 mL of water or test solutions in Petri dishes and incubated at 25°C in the dark. For hormone treatments seeds were imbibed in the presence of 100 µM GA₄₊₇ and/or 100 µM (±)ABA.

In some cases, seeds were dissected into the micropylar tip and the remainder of the seed as previously described (Nonogaki, H. *et al.*, *Physiol Plant* 85:167-172 (1992)), and the embryonic tissues were removed from each part using forceps. The embryo-less micropylar tip and the embryo-less remainder of the seed were denoted as the endosperm cap and lateral endosperm, respectively.

Isolation of cDNAs

A cDNA library was constructed using mRNA from *gib-1* seeds imbibed in 100 µM GA₄₊₇ for 24 h using a λZAP Express cDNA Synthesis Kit (Stratagene, La Jolla, California) according to the manufacturer's instructions. The cDNA library was 5 screened by hybridization of nitrocellulose filter plaque replicas with a partial length (0.9 kb) cDNA of the postgerminative mannanase cDNA (Bewley et al. 1997) labeled with ECL labeling reagents (Amersham Pharmacia Biotech, Inc., Piscataway, NJ). Hybridization was done at 42°C overnight after 1 h prehybridization at the same 10 temperature. Following hybridization, the membranes were washed twice for 20 min each at 42°C with 6 M urea, 0.5% (w/v) SDS, 0.5 X SSC and twice for 5 min each at room temperature with 2X SSC. Independent inserts in the library vector pBK-CMV 15 were sequenced by the Advanced Plant Genetics Facility, University of California, Davis. Sequence comparisons were made using DNASTar software (DNASTar Inc., Madison, WI). Signal peptide prediction was performed using the Signal IP V1.1 server (www.cbs.dtu.dk/services/SignalIP).

Expression of Recombinant Protein in Escherichia coli

The coding region (without the signal peptide) of the mannanase cDNA (nucleotides 83 to 1264 of SEQ ID NO: 11) was amplified by PCR using *Bam*HI site-linked forward primer (5'-CGGGATCCTGTGAAGCTAGGGTT-3') and *Xba*I site-linked reverse primer (5'-CGTCTAGACTAAATCTAACCAAATG-3'). The product 20 was digested with *Bam*HI and *Xba*I and ligated into the *Bam*HI and *Xba*I sites of the maltose-binding protein expression vector pMAL-c2 (New England Biolabs, Inc.). The empty vector and the vector containing insertion were transformed into competent cells of 25 a proteinase-deficient strain (BL21) of *E. coli* and the resulting transformant cells were selected using blue-white screening with IPTG-Xgal plates. After incubation of a 1% overnight culture for 4 h at 37°C, protein expression was induced by addition of IPTG to a final concentration of 2 mM and further incubation at 37°C for 2 h. The bacterial cells 30 were harvested by centrifugation at 6,000 g and the pellet was dissolved in sonication buffer (50 mM sodium phosphate buffer pH 8.0 containing 0.3 M NaCl, 1 mg/mL lysozyme [Boehringer Mannheim]). After overnight freezing at -20°C, the bacterial lysates were thawed and centrifuged at 10,000 g for 10 min and the supernatants were collected. Expressed proteins were visualized by SDS-PAGE of the supernatants (crude extracts) of induced or uninduced bacterial cultures with or without insertion. For

purification of the fusion protein, the supernatant of an induced bacterial culture with insertion was applied to a maltose-binding protein affinity column (amylose resin, New England Biolabs, Inc.). The fusion protein was eluted from the column with 10 mM maltose. The fractions containing the fusion protein were examined by SDS-PAGE, 5 mixed and dialyzed against 10 mM Tris-HCl pH 7.5 overnight at 4°C. The dialysate was stored at -80°C.

PAGE and Immunoblotting

Proteins were separated by SDS-PAGE using 10% (w/v) acrylamide gels 10 according to Laemmli (1970). Native PAGE was done in 7.5% gels according to Davis (1964) except that ammonium peroxydisulfate was used in place of riboflavin in the stacking gel. After electrophoresis, proteins were transferred to PVDF membranes using semidry blotter (TRAN-BLOT SD, BioRad Laboratories, Hercules, CA) and blocked 15 with 5% (w/v) nonfat milk in PBS containing 0.5% (v/v) Tween 20. Anti-maltose binding protein antibody (New England Biolabs) and anti-M3 mannanase antibody were used for immunoblotting at 1:5000 dilution. Bound antibody was detected using horseradish peroxidase-conjugated goat anti-rabbit immunoglobulin (Sigma Immunochemicals, St. Louis, MO). The bands were visualized on X-ray film after the reaction with a chemiluminescence reagent (Renaissance; DuPont NEN Products, Boston, 20 MA).

Endo- β -Mannanase Extraction and Assays

Endo- β -mannanase activity was extracted from tomato seeds or seed parts by homogenizing the tissues in 50 mM K-phosphate buffer, pH 6.8 with a mortar and 25 pestle. The homogenate was centrifuged at 10,000 g for 5 min and the supernatant was used as the enzyme solution. The endo- β -mannanase activity was assayed by the modified gel diffusion method (Still, D.W. *et al.*, *Plant Physiol* 113:21-29 (1997)). Agarose (0.8% [w/v]) plates containing 0.05% (w/v) locust bean galactomannan (Sigma, St. Louis, MO) were solidified and wells were formed on the plates by scoring with a 3- 30 mm cork borer and removing the plug by suction. The extracts (10 μ L) from tomato seed parts or purified recombinant protein solution (10 μ L) were applied to the wells and the plates were incubated at 25°C for 24 h. After incubation, the agarose gel plates were stained by 0.5% (w/v) Congo red dye (Sigma) as described previously (Still, D.W. *et al.*,

Plant Physiol 113:21-29 (1997)). The hydrolyzed areas were visible as clear circles on a dark background. The diameter of the hydrolyzed area is logarithmically related to the enzyme activity. This agarose gel method was also used for activity staining of native PAGE gels. After electrophoresis, a native gel was overlaid on top of the substrate-containing agarose gel and incubated at 25°C for 1 h. The activity band was visualized as an transparent zone on the substrate gel after staining as described above.

DNA Extraction and Southern Hybridization

Genomic DNA was isolated from young tomato leaves (cv. Moneymaker) as described by Murray and Thompson (1980). Genomic DNA (10 µg) was digested with the restriction enzymes *Bam*HI, *Xba*I and *Xho*I (New England Biolabs, Beverly, MA), separated on a 1.0% (w/v) agarose gel, and transferred to positively charged membranes (Hybond-N⁺, Amersham Pharmacia Biotech Inc.). Prehybridization, hybridization, washing, and detection were performed as described for cDNA library screening. For preparing the DNA probes, the vectors containing the full-length mannanase cDNAs were digested with *Bam*HI and *Xho*I, and the gel-purified insertion was used to make ECL-labeled probes (Amersham Pharmacia Biotech Inc.).

RNA Extraction and Northern Hybridization

Total RNA was extracted from seed parts (endosperm cap, lateral endosperm or whole embryo) of germinating or germinated tomato seeds using a standard phenol extraction method (Sambrook, J. et al., *Molecular Cloning. A Laboratory Manual*, Second Edition. N Ford, C Nolan, M Ferguson eds, Cold Spring Harbor Laboratory Press, Cold Spring Harbor (1989)). Total RNA (2 to 10 µg) was subjected to electrophoresis on 1.3% (w/v) agarose gels containing 7% (v/v) formaldehyde, transferred to a neutral membrane (Hybond-N, Amersham Pharmacia Biotech) and UV-crosslinked. RNA probes were prepared using a DIG-labeled dNTP mixture (Boehringer Mannheim, Inc., Indianapolis, IN) Hybridization was routinely done at 60°C overnight after 1 h prehybridization at the same temperature. The membranes were washed once for 25 min with 2X SSC, 0.1% (w/v) SDS at 70°C and twice for 25 min with 0.2X SSC, 0.1% (w/v) SDS at 70°C. The membranes were then blocked for 1 h with 5% (w/v) nonfat milk in 0.1 M maleic acid buffer pH 7.5 containing 0.15 M NaCl, 0.3% (v/v) Tween 20 (buffer A) and incubated with alkaline phosphatase conjugated anti-DIG antibody (1:15,000 dilution) for 1 h at 25°C. After washing with buffer A, the membranes were subjected to

chemiluminescence detection. The signal was detected on X-ray film after 5 to 20 min exposures. When the signals on two different membranes hybridized with different probes were compared, those membranes were exposed together on the same X-ray film for the same time.

5

RESULTS

Isolation of the Germinative Mannanase cDNA

As the expression of mannanase in tomato seeds before radicle protrusion is induced by GA (Groot, S.P.C. et al., *Planta* 174:500-504 (1988)), a cDNA library of 10 GA-induced *gib-1* seeds was screened with a partial length (0.9 kb) cDNA of the postgerminative mannanase (Bewley et al., *Plant Cell* 9:1055-1066 (1997)). Four positive clones isolated from cDNA library screening had sequences similar to that of the postgerminative mannanase cDNA. The longest cDNA insert of the putative mannanase was rescued into pBK-CMV vector, then subcloned into pBluescript II KS, and 15 designated *LeMAN2* (*Lycopersicon esculentum* mannanase 2). In this paper, we designate the cDNA of the postgerminative mannanase as *LeMAN1* because it was the first mannanase cDNA isolated from tomato seeds.

The 1481 bp *LeMAN2* cDNA (nucleotide sequence deposited in Genbank as Accession No. AF184238) contained an open reading frame encoding a protein of 415 20 amino acids. A putative signal peptide sequence of 22 amino acids was identified at the amino terminus of the protein. The mature protein of 393 amino acids was encoded from Cys residue at nucleotide 83 to Ile at nucleotide 1259, with a predicted molecular weight of 44,379 and pI of pH 5.7. The predicted amino acid sequence of the protein encoded by *LeMAN2* was compared with the postgerminative mannanase protein encoded by 25 *LeMAN1*, with an Arabidopsis genomic sequence and with two fungal mannanases. Overall amino acid sequence similarity between *LeMAN1* and *LeMAN2* was 77%. While overall sequence similarity was high, *LeMAN2* contained additional amino acids 30 compared with *LeMAN1* protein (e.g., amino acids 26-35, 132-159), accounting for the greater predicted size of *LeMAN2* (44 kDa versus 38 kDa for *LeMAN1*). Potential catalytic sites and a potential N-glycosylation site (Asn-Gly-Ser) (amino acids 50-52) that have been identified in the postgerminative mannanase were also present in *LeMAN2*. protein. The amino acid sequence similarity between the tomato mannanases (*LeMAN1* and *LeMAN2*) and fungal mannanases (*Aspergillus aculeatus* and *Trichoderma reesei* [Accession number AAA34208]) was approximately 30%. Interestingly, the amino acid

sequences in LeMAN2 that were deleted in the LeMAN1 protein still showed identity to the fungal mannanase proteins.

Expression of the Protein Encoded by LeMAN2 cDNA

To confirm that the *LeMAN2* cDNA encodes endo- β -mannanase protein,
5 the cDNA was inserted into a maltose-binding protein overexpression vector and transformed into *E. coli*. When the transformed cells were induced for protein expression by adding IPTG, a strong intensity band with an apparent molecular mass of 90 kDa was observed in the bacterial lysates, close to the predicted size of the fusion protein (maltose-binding protein [43 kDa] plus LeMAN2 mannanase [44 kDa]). This protein band was
10 absent in the uninduced cells and in both induced and uninduced cells containing the empty vector. The putative fusion protein band was recognized by both anti-maltose-binding protein antibody and antibody to one of the postgerminative mannanases (anti-M3 mannanase antibody). These results confirm that the overexpressed protein contains the maltose-binding::mannanase fusion protein. Extracts of the induced bacterial cells
15 containing the *LeMAN2* insert showed endo- β -mannanase activity, which could not be detected in bacterial cells that contained empty pMAL vector, indicating that the overexpressed recombinant protein was an active form of mannanase. When the fusion protein was affinity purified to homogeneity by using a maltose-binding protein affinity resin, the fractions containing the fusion protein showed high mannanase activity and
20 were recognized by the anti-mannanase antibody. The gel diffusion assay method for mannohydrolase activity is specific for endo-type enzymes, so *LeMAN2* clearly encodes an endo- β -mannanase.

Southern Hybridization

The hybridization patterns of the *LeMAN1* and *LeMAN2* cDNAs with
25 tomato genomic DNA were compared using Southern hybridization. Both cDNAs hybridized to the same sets of DNA fragments, confirming that multiple mannanase genes are present in the tomato genome. However, some DNA fragments hybridized more strongly to the *LeMAN1* cDNA, while other bands showed a stronger signal with the
30 *LeMAN2* cDNA. This supports the sequence data indicating that different genes encode the germinative and postgerminative mannanases.

Expression of LeMAN1 and LeMAN2 mRNA in tomato seeds

Since *LeMAN2* was isolated from a cDNA library prepared from tomato seeds prior to radicle emergence, the *LeMAN2* protein is likely to be a germinative mannanase. Only the M α germinative mannanase specific to the endosperm cap is present in the endosperm at this time (Nonogaki, H. et al., *Plant Physiol.* 110:555-559 (1996)), but two embryo-specific mannanases are also present in germinating tomato seeds (Nonogaki et al., *Physiol. Plant.* 102:236-242 (1998)). To investigate in which tissue(s) of imbibed seeds the *LeMAN2* mRNA is expressed, RNA gel blot analyses were performed. When the total RNA from dissected seed parts (endosperm cap, lateral endosperm and embryo) from wild type tomato seeds imbibed in water for 24 h was hybridized with a full-length *LeMAN2* RNA probe, the transcript was detected only from the endosperm cap, indicating that *LeMAN2* mRNA is specifically expressed in this tissue. The endosperm cap tissue also contained high mannanase activity, while little or no activity was detected from the lateral endosperm. Although some mannanase activity was present in the embryo as well, no hybridization between the *LeMAN2* and embryonic mRNA was detected. Since the northern hybridization was performed under relatively high stringency conditions (70°C washing), hybridization at low stringency (55°C) was also examined. However, even at low stringency, no signal could be detected in the embryo.

To compare the tissue specificity of expression of *LeMAN2* (see above) with that of *LeMAN1*, total RNA was extracted from the endosperm caps of seeds prior to radicle emergence and from the lateral endosperms of germinated seedlings at different stages of development. To directly compare the hybridization patterns, the same sets of RNA samples were loaded on the same gel, transferred to the same membrane, processed using the same anti-DIG-antibody solution following hybridization to the different probes, and exposed to the same X-ray film. The patterns of hybridization by *LeMAN1* and *LeMAN2* riboprobes to these RNA samples were completely different. When the *LeMAN1* probe was used, hybridization was detected specifically in the RNA samples from postgerminative lateral endosperms after radicle growth had begun, although a faint band could also be seen in the endosperm caps prior to radicle emergence after a longer exposure to the X-ray film. On the other hand, when *LeMAN2* was used as a probe, a strong signal of the transcript was detected in the RNA sample from the endosperm cap of seeds prior to radicle emergence, and only faint signals were detected at post-emergence

stages. Thus, under the conditions used, there is little cross-hybridization of riboprobes prepared from each cDNA. The results demonstrate that expression of *LeMAN2* is specific to the endospenn cap prior to radicle emergence and that *LeMAN1* expression is localized to the lateral endosperm after radicle emergence.

Given the timing and location of its expression, it is likely that *LeMAN2* is involved in cell wall hydrolysis associated with endosperm cap weakening prior to radicle protrusion. To examine the correlation between the expression of *LeMAN2* and germination, the abundance of *LeMAN2* message in the endosperm caps of imbibed wild type tomato seeds was analyzed. It is known that there is no detectable mannanase activity in dry tomato seeds and that activity begins to increase 6 to 12 h after imbibition (Groot, S.P.C. *et al.*, *Planta* 174:500-504 (1988)). *LeMAN2* transcript was present in the endosperm cap within 12 h of imbibition and increased markedly by 24 h before declining slightly by 36 h. The timing of *LeMAN2* expression corresponded to the appearance of mannanase activity in the same tissue, although the peak of mRNA accumulation occurred earlier than the maximum activity, as would be expected. Radicle protrusion of wild type seeds was first observed 40 to 48 h after imbibition (data not shown), well after the increase in *LeMAN2* message and mannanase activity.

Hormonal regulation of the expression of *LeMAN2* in the wild type- and *gib-1* tomato seeds was also analyzed. While wild type seeds can germinate in water, germination of *gib-1* mutant tomato seeds is dependent on application of exogenous GA. *LeMAN2* mRNA could not be detected in the endosperm caps of *gib-1* mutant seeds incubated in water and mannanase activity was barely detectable in this condition. In contrast, both *LeMAN2* mRNA expression and mannanase activity were induced in the endosperm caps of *gib-1* seeds imbibed in GA₄₊₇. The expression of *LeMAN2* mRNA and mannanase activity in the endosperm caps of wild type seeds in water or *gib-1* seeds in GA was not inhibited by 100 µM ABA, although radicle protrusion did not occur in either case (data not shown). These findings are consistent with previous observations that germinative mannanase activity in the endosperm cap is insensitive to ABA (Toorop, P.E. *et al.*, *Planta* 200:153-158 (1996); Dahal P. *et al.*, *Plant Physiol* 113:1243-1252 (1997); Still, D.W. *et al.*, *Plant Physiol* 113:13-20 (1997)).

Example 4

This example describes cloning of a subunit of the vacuolar proton-translocating ATPase (LVA-P1, SEQ ID NO: 15).

Tomato (*Lycopersicon esculentum* Mill.) seeds were harvested from field-grown wild type MM (cv. Moneymaker) plants or from homozygous gibberellin-deficient (*gib-1*) mutant plants grown in a glasshouse (seeds originally obtained from Dr. Cees Karssen, Wageningen Agricultural University, The Netherlands). Plant culture and seed extraction were as described previously (Ni, B.R. *et al.*, *Plant Physiol* 101:607-617 (1993)). For germination, approximately 200 seeds were incubated at 25°C in the dark in 9 x 100 mm diameter Petri dishes on top of two layers of blotter paper moistened with 12 mL of either distilled, deionized water or 100 µM GA₄₊₇ (Abbott Laboratories, Chicago, IL, USA)

Differential cDNA Display Analysis

Using DCD (Liang, P. *et al.*, *Science* 257:967-971 (1992)), mRNA from the radicle tips or endosperm caps of *gib-1* mutant seeds imbibed in water (which do not germinate) were compared with mRNA from the same tissues of *gib-1* seeds imbibed in 100 µM GA₄₊₇ (in which radicle emergence begins at 36 to 40 h and is completed by most seeds within 60 h). Seeds in both the presence and absence of GA₄₊₇ were sampled at 40 h, excluding any seeds from which the radicle had emerged. The micropylar regions were excised from 100 seeds in each treatment and separated into endosperm caps and radicle tips. Tissues were frozen immediately in LN₂ and stored at -80°C. Frozen radicle tips or endosperm caps were pulverized in LN₂ and RNA was extracted and purified by the phenol/SDS method (Ausubel F.M. *et al.*, *Current Protocols in Molecular Biology*, Wiley-Interscience, New York (1987)). Prior to use, aliquots of RNA were incubated with DNase I for 1 h at 37°C in digestion buffer (40 mM Tris-HCl pH 7.5, 6 mM MgCl₂, 20 U RNasin and 10 mM NaCl) followed by extraction with 1 volume phenol/chloroform/isoamyl alcohol (25:24:1). The RNA was precipitated in ethanol and dissolved in the original volume of 2 mM EDTA.

Nine 3' anchor primers were synthesized as 5'-T₁₂MM-3', where M is A, G or C. Anchor primers were then combined at equal concentration to give three sets: T₁₂MA, T₁₂MG, and T₁₂MC. These three anchor pools were used in reverse transcription reactions (Sambrook, J. *et al.*, *Molecular Cloning. A Laboratory Manual, Second Edition*. N Ford, C Nolan, M Ferguson eds, Cold Spring Harbor Laboratory Press, Cold Spring Harbor (1989)) and then in conjunction with eight specific ten-mers (A₀₁-A₀₈, Operon

Technologies Inc., Alameda, CA, USA) in the differential display PCR reactions (Liang, P. *et al.*, *Science* 257:967-971 (1992)). Two microliters from the RT reaction were used in each of the subsequent 20 µL differential display reactions (10 mM Tris-HCl pH 8.3, 50 mM KCl, 1.5 mM MgCl₂, 2.5 µM dNTPs, 1 µM of each anchor primer, 0.2 µM 10-mer primer, 1 U AmpliTaq (Perkin-Elmer Cetus, Norwalk, CT, USA), 4 µCi [α -³⁵S]-dATP (1200 Ci/mmol). Cycling conditions were 30 s at 94°C, 2 mm annealing at 40°C and 30 s extension at 72°C for 40 cycles in a Perkin-Elmer Cetus 480 Thermal Cycler. PCR reactions were loaded onto a 40 cm x 40 cm x 0.4 mm 6% native polyacrylamide gel and electrophoresed at 40 watts.

Selected cDNA fragment bands exhibiting differential amplification were cut from the DCD gels and recovered by boiling for 30 mm in 50 µL modified TE (10 mM Tris-HCl pH 7.5, 0.1 mM EDTA). The cDNAs were precipitated in ethanol in the presence of 10 µg linear acrylamide as a carrier (Gaillard, C *et al.*, *Nuc Acids Res* 18:378 (1990)) and dissolved in 20 µL TE. Five microliters of each isolated cDNA was re-amplified using the same conditions as the first amplification except that the (now unlabelled) dNTP concentration was increased to 20 µM.

Fragments re-amplified from DCD gels were tested for differential expression using the reverse-northern technique. Four identical blots of electrophoresced DNA fragments were probed with labeled cDNA products from reverse transcription reactions using 10 µg total RNA from endosperm caps or radicle tips of seeds imbibed for 40 h in water or 100 µM GA₄₊₇ as template and labeled using 200 µM dNTP and 100 µCi (α -³²P) dCTP (3000 Ci/mmol) per 50 µL reaction.

PCP fragments selected on the basis of the reverse-northern results were ligated into the TA cloning vector pCRII (Invitrogen Corp., San Diego, CA, USA) and the resulting plasmids electroporated (Cooley, M.B. *et al.*, *J Bacteriol* 173:2608-2626 (1991)) into competent *E. coli* JM109 cells (Stratagene, La Jolla, CA, USA) using the Gene Pulser Electroporator (Bio-Rad Laboratories, Hercules, CA, USA). The cloned DCD fragments were sequenced at the UC Davis Advanced Plant Genetics Facility on an ABI Prism 377 DNA Sequencer (ABI, Perkin-Elmer, Foster City, CA).

30

Northern Analyses

Total RNA was extracted as described above from intact wild type MM or *gib-1* mutant seeds or isolated endosperm caps, radicle tips, and the rest of the seed

incubated in water or in 100 μM GA₄₊₇ for the indicated times at 25°C. Additionally, total RNA was isolated from the flowers, leaves and roots of mature MM tomato plants. Total RNA (2 to 20 μg per lane) was electrophoresed (Sambrook, J. et al., *Molecular Cloning. A Laboratory Manual, Second Edition.* N Ford, C Nolan, M Ferguson eds, Cold Spring Harbor Laboratory Press, Cold Spring Harbor (1989)), transferred onto positively charged nylon membranes and UV crosslinked at 120,000 $\mu\text{J cm}^{-2}$ on a FB-UVXL-1000 Stratalinker (Fisher Scientific, Santa Clara, CA, USA). Hybridization was detected using DIG-labeled RNA probes (Boehringer Mannheim Corp., Indianapolis, IN, USA) synthesized by either Sp6 (Ambion Inc., Austin, TX, USA) or T7 (Pharmacia Biotechnology, Inc., Alameda, CA, USA) RNA polymerase. Detection of DIG-labeled probes was performed according to instructions in the Genius System (Boehringer Mannheim Corporation. *Genius System User's Guide for Membrane Hybridization. Version 3.0.* (1995)) using CSPD (disodium 3-(4-methoxyspiro{1,2-dioxetane-3,2'-(5'-chloro)tricyclo[3.3.1.1^{3,7}]decan}-4-yl) phenyl phosphate) as substrate.

15

cDNA Library Screening

Transcripts hybridizing to PCR fragment *G21*, which were differentially expressed in seeds in response to GA, were shown by northern analysis to be expressed in tomato roots as well as in seeds. Hence, 5×10^5 recombinants from a tomato root cDNA library in the plasmid pCGN1703 (Ewing, N.N. et al., *Plant Physiol* 94:1874-1881 (1990)) were screened with *G21* labeled in a random priming reaction with α -(³²P)-dATP. cDNA from hybridizing recombinants was recovered from the vector by restriction digestion with *Sma*I and was subcloned into the *Sma*I site in the polycloning region of pBSII KS (Stratagene, La Jolla, CA). The full-length cDNA hybridizing to *G21*, subsequently termed *LVA-P1*, was sequenced at the Center for Engineering Plants for Resistance Against Pathogens (CEPRAP), University of California, Davis.

20

Protein Extraction and Western Blotting

Two grams of *gib-1* mutant tomato seeds imbibed in water or 100 μM GA₄₊₇ for 36 h were homogenized for 1 mm in an Ultra-Turrax T25 (Janke and Kunkel IKA Labortechnik, Staufen, Germany) at high speed in 5 mL of extraction buffer (70 mM Tris pH 8.0, 250 mM sucrose, 3 mM EDTA, 0.5% PVP-40, 0.1% BSA, 4 mM DTT). The slurry was filtered through cheesecloth and the filtrate was centrifuged at 12,000 *g* for 15

mm at 4°C and the supernatant again centrifuged at 113,000 g for 30 minutes at 4°C. The resulting pellet was resuspended in buffer (10 mM Tris/Mes pH 7; 250 mM sucrose, 1 mM DTT) and electrophoresced on 12% SDS-polyacrylamide (20 µg of total protein per lane). The proteins were transferred to nitrocellulose and detected as described below for 5 western tissue prints.

Northern and Western Tissue Printing Protocols

For tissue printing (Reid, P.D. et al., *Tissue Printing: Tools for the Study of Anatomy, Histochemistry, and Gene Expression*. Academic Press, New York (1992)), 10 tomato seeds were sliced longitudinally with a fresh razor blade. Each half-seed was pressed firmly in identical positions on separate nitrocellulose membranes for exactly 60 s (northerns) or 20 s (westerns), then removed with forceps, providing two mirror-image prints of the same seed. The entire procedure was performed using powder-free gloves.

For northern prints, the membranes were UV crosslinked and treated with 15 10 U RNase-free DNase I (Pharmacia Biotechnology, Inc., Alameda, CA, USA) in 10 mL of digestion buffer for 1 h at 37°C in a roller tube. Subsequent steps in prehybridization, hybridization, washing and detection were as described for the DIG-labeled northerns (see above) except that the signal was detected using CDP-Star (disodium 4-chloro-3-(4-methoxyspiro{1,2-dioxetane-3,2'-(5'-chloro)tricyclo[3.3.1^{3,7}]decan}-4-yl) phenyl 20 phosphate; Boehringer Mannheim) in conjunction with enhancer for chemiluminescent detection on nitrocellulose membranes (Tropix, Inc., Bedford, MA, USA). One membrane was hybridized with an antisense probe to detect the target mRNA while its mirror-image print was hybridized with a sense probe to control for nonspecific binding. Both membranes were then stripped and rehybridized with antisense and sense cDNAs 25 complementary to a constitutively expressed mRNA coding for a ribosomal protein (*G46*) as a control for RNA bound to the membrane.

For western prints, the printed membranes were blocked using 1% BSA in 1x TBS, washed four times for 5 min each in 1x TBS-Tween (Sambrook, J. et al., *Molecular Cloning. A Laboratory Manual, Second Edition*. N Ford, C Nolan, M 30 Ferguson eds, Cold Spring Harbor Laboratory Press, Cold Spring Harbor (1989)) and then incubated with primary antibody diluted 1:4000 in 1% BSA in 1x TBS. Rabbit polyclonal antibodies specific for either subunit A (68 kD) or subunit B (57 kD) of V-ATPase from mung bean (*Vigna radiata*) were provided by Dr. Masayoshi Maeshima (Matsuura-Endo, C. et al., *Plant Physiol* 100:718-722 (1992)). The membranes were

subsequently washed in 1x TBS-Tween and incubated with a 1:10,000 dilution of goat anti-rabbit IgG alkaline phosphatase-conjugated secondary antibody. After washing, the membranes were developed using NBT-BCIP (nitroblue tetrazolium chloride and 5-bromo-4-chloro-3-indolyl-phosphate; Boehringer Mannheim).

5

RESULTS

Cloning of a V-ATPase Subunit c Gene from Germinating Tomato Seeds

Pools of total RNA from *gib-1* tomato radicle tips and endosperm caps were isolated separately after imbibition in either water or 100 µM GA₄₊₇ for 40 h, or just prior to the initiation of radicle emergence in the presence of GA₄₊₇ (radicle emergence does not occur in water). cDNA fragments identified by DCD as being differentially expressed were confirmed by northern analysis. One such fragment (*G21*) identified a transcript that increased in abundance in both endosperm caps and radicle tips in the presence of GA₄₊₇ and was present in untreated roots (data not shown). This fragment was used to recover a homologous full-length cDNA from a tomato root cDNA library. The deduced amino acid sequence is highly homologous to that of the 16 kD hydrophobic subunit c that forms the membrane-spanning, proton conductance pathway of plant vacuolar H⁺-translocating ATPases (Stevens, T.H. et al., *Annu Rev Cell Dev Biol* 13:779-808 (1997); Sze, H. et al., *J Bioenerg Biomem* 24:371-381 (1995)). For example, the tomato cDNA sequence showed the following percentage amino acid identities (nucleotide identities in parentheses) to other V-ATPase c subunits: 98.2% (84%) to *Arabidopsis thaliana* (Accession No. L44581; Perera, I.Y. et al., *Plant Mol Biol* 29:227-244 (1995)); 98.2% (85%) to *Gossypium hirsutum* (U13669; Hasenfratz, M-P. et al., *Plant Physiol* 108:1395-1404 (1995)); and 95.8% (79%) to *Avena sativa* (M73232; Lai, S. et al., *J Biol Chem* 266:16078-16084 (1991)). We have therefore named this cDNA *LVA-P1* (*Lycopersicon* Vacuolar ATPase-Proteolipid 1), in analogy with *AVA-P1* and related genes in *Arabidopsis* (Perera, I.Y. et al., *Plant Mol Biol* 29:227-244 (1995)). The highly conserved amino acid sequence among species (including over 60% identity to corresponding mouse [M64298] and yeast [LO7105] genes) and the >98% amino acid sequence identity of *LVA-P1* to other dicot V-ATPase subunit c genes leave little doubt that *LVA-P1* is a tomato homolog of this gene.

LVA-P1 Expression Patterns During Germination

To determine the expression pattern of *LVA-P1*, total RNA was extracted from endosperm caps, radicle tips, and the rest of *gib-1* seeds (includes the lateral endosperm and most of the embryo) after imbibition for 1, 12, 24, and 40 h in water or in 5 $100 \mu\text{M}$ GA₄₊₇ and hybridized with full-length *LVA-P1* cDNA. In the presence of GA₄₊₇, *LVA-P1* transcript accumulated preferentially in the micropylar region of the seed (endosperm cap and radicle tip) within 12 h of imbibition. Subsequently, *LVA-P1* mRNA abundance in the endosperm caps declined by 40 h but remained high in the radicle tips. 10 Approximately 30% of the seeds had completed radicle emergence by 40 h, although RNA was extracted only from ungerminated seeds. Transcripts hybridizing to *LVA-P1* were also detected in wild type MM flowers, leaves and roots.

Previous results have documented that a wide range of enzyme activity can exist among individual seeds, even in homozygous inbred tomato lines (Still, D.W. *et al.*, *Plant Physiol* 113:21-29 (1997)). Thus, mRNA extracted from pooled seed samples may 15 not accurately reflect individual seed responses to GA. We therefore utilized tissue printing to assay mRNA abundance on a single-seed basis. Individual seeds were bisected after various times of imbibition and each mirror-image half was printed in an ordered array on separate nitrocellulose membranes. The membranes were treated with DNase, then hybridized with riboprobes made from either the antisense or the sense 20 strand of the cDNA. This tissue printing method was specific, with little or no hybridization detected with riboprobes made from the sense strands of *LVA-P1* or *G46*. Hybridization of the antisense riboprobe to the constitutive *G46* mRNA, however, showed that approximately equal amounts of RNA were bound to the membrane by each 25 seed. Hybridization of the antisense *LVA-P1* riboprobe revealed an increase in *LVA-P1* mRNA abundance only in the presence of GA₄₊₇. Furthermore, in most seeds, *LVA-P1* mRNA was most abundant in the micropylar region, in agreement with the pattern inferred from the northern blots of pooled samples.

As *LVA-P1* expression in *gib-1* seeds was dependent upon GA₄₊₇, transcript abundance was investigated in wild type MM seeds, which do not require 30 additional GA for germination. *LVA-P1* mRNA was present during seed development but declined in seeds from mature green and breaker stages of fruit development, before increasing again in seeds from ripe fruit (0 h of imbibition). Following imbibition, whole seeds were sampled every 12 h and separated into germinated and ungerminated seeds at 48 and 60 h. *LVA-P1* mRNA content changed relatively little prior to radicle emergence

and remained abundant in germinated seeds. As was observed for *gib-1* seeds in the presence of GA₄₊₇, *LVA-P1* mRNA was most abundant in the micropylar tissues. Imbibition of MM seeds in 100 µM GA₄₊₇ had no additional effect on mRNA abundance, indicating that the endogenous GA content of wild type seeds is sufficient to saturate the response.

GA-dependent Expression of V-ATPase Protein Subunits in gib-1 Seeds

As GA stimulated the expression of *LVA-P1* mRNA in *gib-1* seeds, we tested whether other protein subunits associated with the V-ATPase complex were induced as well. The membrane-spanning hydrophobic subunit c coded by *LVA-P1* is difficult to extract and detect with antibodies, so western blots of proteins from *gib-1* seeds imbibed in either water or GA₄₊₇ were performed using antibodies specific to the major 68 kDa catalytic nucleotide-binding subunit A and the 57 kDa noncatalytic nucleotide-binding subunit B of mung bean V-ATPase (Matsuura-Endo, C. et al., *Plant Physiol* 100:718-722 (1992); Maeshima, M. et al., *Plant Physiol* 106:61-69 (1994)). The V-ATPase complex requires the membrane-spanning subunit c for assembly of the V₀ membrane sector, to which the cytoplasmic V₁ sector containing subunits A and B is attached (Sze, H. et al., *J Bioenerg Biomem* 24:371-381 (1995)). Thus, detection of both A and B subunits is likely to be a good indicator of the presence of the holoenzyme, including the subunit c protein. Protein bands of the expected size increased in intensity in extracts from *gib-1* seeds that had been imbibed in GA₄₊₇. No other proteins were detected, so the antibody for the 57 kDa subunit B was used with tissue prints to determine the localization of the V-ATPase within the seed. In agreement with the expression pattern of *LVA-P1* mRNA, the subunit B protein was most abundant in the micropylar region of GA-treated *gib-1* seeds, particularly in the endosperm cap, while seeds imbibed in water exhibited only background staining.

It is understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and scope of the appended claims. All publications, patents, and patent applications cited herein are hereby incorporated by reference in their entirety for all purposes.

WHAT IS CLAIMED IS:

- 1 1. An isolated nucleic acid molecule comprising a polynucleotide
2 sequence encoding a polypeptide that modulates seed germination, wherein the
3 polypeptide comprises an amino acid sequence that has at least about 70% identity to
4 SEQ ID NOs: 4, 6, 8, 10, 14, 18, 20, and 22.
- 1 2. The nucleic acid molecule of claim 1, wherein the polynucleotide
2 sequence specifically hybridizes to SEQ ID NO: 3, 5, 7, 9, 13, 17, 19, 21 or their
3 complements.
- 1 3. The nucleic acid molecule of claim 1, wherein the polypeptide is as
2 shown in SEQ ID NO: 4, 6, 8, 10, 14, 18, 20, and 22.
- 1 4. The nucleic acid molecule of claim 1, wherein the polynucleotide
2 sequence is as shown in SEQ ID NO: 3, 5, 7, 9, 13, 17, 19, or 21.
- 1 5. A recombinant expression vector comprising the polynucleotide
2 sequence of claim 1.
- 1 6. The recombinant expression vector of claim 5, further comprising a
2 promoter operably linked to the polynucleotide sequence.
- 1 7. The recombinant expression vector of claim 6, wherein the
2 promoter is a constitutive promoter.
- 1 8. The recombinant expression vector of claim 6, wherein the
2 promoter is an inducible promoter.
- 1 9. A host cell transformed with the recombinant expression vector of
2 claim 5.
- 1 10. A transgenic plant comprising a recombinant expression cassette
2 comprising a promoter operably linked to the polynucleotide sequence of claim 1.
- 1 11. The transgenic plant of claim 10, wherein the plant is tomato.
- 1 12. A method of modulating seed germination in a plant, the method
2 comprising

3 a) introducing into the plant a recombinant expression vector comprising a
4 promoter operably linked to a polynucleotide sequence encoding a polypeptide that
5 modulates seed germination, wherein the polypeptide comprises an amino acid sequence
6 that has at least about 70% identity to SEQ ID NOs: 2, 4, 6, 8, 10, 14, 18, 20, and 22; and
7 b) selecting a plant with modulated seed germination.

1 13. The method of claim 12, wherein the polynucleotide specifically
2 hybridizes to SEQ ID NO: 1, 3, 5, 7, 9, 13, 17, 19, 21 or their complements.

1 14. — The method of claim 12, wherein the polypeptide is as shown in
2 SEQ ID NO: 2, 4, 6, 8, 10, 14, 18, 20, and 22.

1 15. An isolated nucleic acid molecule comprising a promoter sequence
2 from SEQ ID NO: 23 operably linked to heterologous nucleic acid sequence.

1 16. An isolated nucleic acid molecule comprising a polynucleotide
2 sequence encoding a polypeptide that modulates seed germination, wherein the
3 polypeptide comprises an amino acid sequence that has 99% or greater identity to SEQ ID
4 NO: 16.

1 17. The nucleic acid molecule of claim 16, wherein the polynucleotide
2 sequence specifically hybridizes to SEQ ID NO: 15 or its complement.

1 18. The nucleic acid molecule of claim 16, wherein the polypeptide is
2 as shown in SEQ ID NO: 16.

1 19. The nucleic acid molecule of claim 1, wherein the polynucleotide
2 sequence is as shown in SEQ ID NO:15.

1 20. A recombinant expression vector comprising the polynucleotide
2 sequence of claim 16.

1 21. The recombinant expression vector of claim 20, further comprising
2 a promoter operably linked to the polynucleotide sequence.

1 22. The recombinant expression vector of claim 20, wherein the
2 promoter is a constitutive promoter.

- 1 23. The recombinant expression vector of claim 20, wherein the
- 2 promoter is an inducible promoter.
- 1 24. A host cell transformed with the recombinant expression vector of
- 2 claim 20.
- 1 25. A transgenic plant comprising a recombinant expression cassette
- 2 comprising a promoter operably linked to the polynucleotide sequence of claim 16.
- 1 26. The transgenic plant of claim 25, wherein the plant is tomato.
- 1 27. A method of modulating seed germination in a plant, the method
- 2 comprising
 - 3 a) introducing into the plant a recombinant expression vector comprising a
 - 4 promoter operably linked to a polynucleotide sequence encoding a polypeptide that
 - 5 modulates seed germination, wherein the polypeptide comprises an amino acid sequence
 - 6 that has 99% or greater identity to SEQ ID NO: 16; and
 - 7 b) selecting a plant with modulated seed germination.
- 1 28. The method of claim 27, wherein the polynucleotide specifically
- 2 hybridizes to SEQ ID NO:15 or its complement.
- 1 29. The method of claim 27, wherein the polypeptide is as shown in
- 2 SEQ ID NO:16.
- 1 30. An isolated nucleic acid molecule comprising a polynucleotide
- 2 sequence encoding a polypeptide that modulates seed germination, wherein the
- 3 polypeptide comprises an amino acid sequence that has 80% or greater identity to SEQ ID
- 4 NO: 12.
- 1 31. The nucleic acid molecule of claim 30, wherein the polynucleotide
- 2 sequence specifically hybridizes to SEQ ID NO: 11 or its complement.
- 1 32. The nucleic acid molecule of claim 30, wherein the polypeptide is
- 2 as shown in SEQ ID NO:12.
- 1 33. The nucleic acid molecule of claim 30, wherein the polynucleotide
- 2 sequence is as shown in SEQ ID NO:12.

1 34. A recombinant expression vector comprising the polynucleotide
2 sequence of claim 30.

1 35. The recombinant expression vector of claim 34, further comprising
2 a promoter operably linked to the polynucleotide sequence.

1 36. The recombinant expression vector of claim 34, wherein the
2 promoter is a constitutive promoter.

1 37. The recombinant expression vector of claim 34, wherein the
2 promoter is an inducible promoter.

1 38. A host cell transformed with the recombinant expression vector of
2 claim 34.

1 39. A transgenic plant comprising a recombinant expression cassette
2 comprising a promoter operably linked to the polynucleotide sequence of claim 30.

1 40. The transgenic plant of claim 39, wherein the plant is tomato.

1 41. A method of modulating seed germination in a plant, the method
2 comprising

3 a) introducing into the plant a recombinant expression vector comprising a
4 promoter operably linked to a polynucleotide sequence encoding a polypeptide that
5 modulates seed germination, wherein the polypeptide comprises an amino acid sequence
6 that has 80% or greater identity to SEQ ID NO:12; and

7 b) selecting a plant with modulated seed germination.

1 42. The method of claim 41, wherein the polynucleotide specifically
2 hybridizes to SEQ ID NO:11 or its complement.

1 43. The method of claim 41, wherein the polypeptide is as shown in
2 SEQ ID NO:12.

SEQ ID NO: 1

LeEXP4 (Expansin) cDNA from tomato seed, 1213 bp

ATTTCTCTCTACTTCTTAACACTGAGCAGTAAGCGAGTGAGTATGAGAAAAATGGCTGCCAATATGATG
CTCTACATTACTATTACTGTTCTCTGTCTCACTGCCGTCAATGCCAGAATCCCCGGCGTTATACCGGCGGACC
ATGGCAAACCGCCCACGCCACCTCTACGGTGGCTCTGACGCATCTGGAACTATGGGTGGAGCTTGTGGATATGGCAATT
TATACAGCCAAGGTTACGGAGTGAATAATGCAGCGTTAACGACAGTGCTATTAACAATGGACTAACGTCGGAGCGTGC
TTTGAATTAAGTGTGATAACGATGGCAAATGGTGTCTCCTGGTAATCCATCCATTCTGTGACGGCGACAAATTTTG
CCCGCCGAACCTCGCTTACCAAACGATGACGGCGGGTGGTGTAAACCTCCTCGTCCTCATTCGATCTGCCATGCCA
TGTTCTCAAAATCGGACTGTACCGTGCCTGAATTGTCCCCGTACATACCGCCAGTACCATGCAGAAAACAAGGAGGA
ATTGATTCACAATAAACGGTTCCGTTACTTCAATTGTTATTGGTAACAAACGTTGCAGGGTGCAGGGGATATACAGAA
GGTCTTAATTAAAGGAACAAACACACAATGGATAGCAATGAGTCGAATTGGGGGAAAATTGGCAAACTAATTCACCTT
TAGTGGGTCAAGCCCTTCTATTGGTTAAAGCTAGTGTACATCGTAGTGTACGAATGTCAACGTGGCACCCCTCTAAT
TGGCAGTTGGACAAACTTGAAGGCAAGAATTCCGGTTAGATCCATAACCCATTCAACTGACCCACCCAAAA
AAACAGAATTACTTAGTATTATACACCAACAAAAACAAGATTTCCTAGACTTTAATTGTTCTCTTTTACTGA
GAAAGTATTGAAGTCTAAGGGATTAGTATCTTTATTATTTGAACTTGAGTATTGTTTTTTTACTA
TTGGTGTGTGAATTGGCTGAAGAGGGTGCACCAACAAAAATGATTAAAGGGAAAAGCATGTAGCCG
CAGCTCTATTGGCATGTTGATGTATTTCTATGAACGAATACCAAAAGATATATAAGTAATTATATAATACTCT
AATGTTTTATAT

SEQ ID NO: 2

LeEXP4 (Expansin) predicted amino acid sequence. 263 amino acids

MRKMAANMMLYITITVLLCFLTAVNARIPGVYTGGPWQTAHATFYGGSDASGTMGGACGYGNLYSQGYGVNNAALSTVLF
NNGLSCGACFELKDNDGKWLPGNPSIFVTATNFCPPNFALPNDDGGWCNPMPRHFDLAMPFLKIGLYRAGIVPVTYR
RVPCRKQGGIRFTINGFRYFNLLLVTNVAGAGDIQKVLIKGNTQWIAMSRNWQQTNSPLVGQALSIRVKASDHRSV
TNVNAPSNWQFGQTATEGKNFRV

SEQ ID NO: 3

LeEXP8 (Expansin) cDNA from tomato seed, 1103 bp

CAAAAAGTAATTCTTAATTTCCACTAAGAAATTAAAATGGCAAATAATGTCATTTAGCATGGGATTTATAATTGG
TTTGTGTACATTTCTCCAGTGCAAATGGTTCTCAGCAGATTCTGGATGGACAAGTGCTCATGCCACATTTATGGT
GGAGCTGATGCTCTGGCACAAATGGGAGGTGCTTGTGGATATGGTAATTGTATTCAACAGGATATGGTACTAGAACAG
CAGCATTGAGTACAGCATTGTTCAATGATGGAGGATCATGTGGTCAATGTTACAAAATAATTGTGATTACAAATTAGA
TCCTCAATGGTCAAAAAGGGAGTATCAGTTACAATTACATCTACAAATTGGTCCACCAAATTATAATCTTCCTAGT
AACAAATGGAGGATGGTGCAACCCCTCACGTCCACATTGATATGGCTCAACCAGCTGGGAAAAAAATTGGCATTACA
AAGGTGGCATTGTTCTGTTTATACAAAAGGGTACCTGCAAGAACATGGTGGAGTTAGATTCAACAAATAATGGAAG
AGACTATTTGAACTAGTGTGGTAAGCAATGTAGGAGGGCAGGATCTGTTGAATCAGTTCAAATCAAAGGCTCAAAC
ACAAATTGGCTAACATGTCAAGAAATTGGGAGCAAGTTGCAATCAAATGCATATCTGATGGACAATCAATATCAT
TCAAAGTTACTACTAGTGTGGTGTACAAAACATTCTAAATATTGTTCCATCTAGTTGGAAATTGGCCAAACATT
TTCAAGCAAAACTCAATTCTGATTGTTAAATGTTGGTGAATACGTTTAATCAGAGGTCTGAGTTATAGGGCAG
CGGCATGCTTACTTTTACATCAAGTAGCCCTCCCCAAAAAAAGAACATCATTCTATTGAGGTATATTGGAGGAGTTG
TTGGTTAAGTAAAGCACATGAATGAACATAGTCAAATAAAGATAGTGA AAAATTAGAAATAACTAAATTATTGGTGTAA
AGTAGGGTAGTTCAAAAATTGTATTTTACTATATGAAAGGATGAATAAGATAAGTTGTTAAAGCTACATC

SEQ ID NO: 4

LeEXP8 (Expansin) predicted amino acid sequence. 257 amino acids

MANNVNLALGFIIGLCTFFSANGFSADSGWTSAHATFYGGADASGTMGGACGYGNLYSTGYGRTAALSTALENDGGSC
GQCYKIIICDYKLDPWCKKGVSVTITSTNFCPPNLYNLPNNNGWCNP PRPHFDMAQPAWEKIGIYKGGIVPVLYKRPVPC
KHGGVRFTINGRDYFELVLVSNVGGAGSVQIKGSNTNWLTMSRNWGASWQSNA LDGQSI SFKVTTSDGVTKTFLNI
VPSSWKFGQTFSKTFQF

SEQ ID NO: 5

LeEXP10 (Expansin) cDNA from tomato seed, 1167 bp

AGTTCCCTAACGCAAACACACAAACATTAAGTTATAAGTGAGCCAAACATCTCTCAATTTCATTTTTTTTAAA
AAAAATGTCAATTGTGGTCTCCATAATTGGTGTGGTATACAATTATAATTGTGCCATGGAAATGAACAAG
GATGGATTGAAGCTCATGCAACTTCTATGGAGGTGGTATGCCGGAACTATGGGTGGAGCATGTGGTTATGGGAAT
TTGTACAGTGAAGGATATGGTACAAACACAGCAGCATTGAGTACAGCACTGTTCAATAATGGTTGAGTTGTGGATCTG
CTTGAGCTAAATGTGTGGGTGATCGAAGTGGTGCCTCCAGGCTCTAGTGGTAAGTGTACTAATTGCCCCAC
CAAATTGCCCCTCCGAATAACGCTGGTGGTGCAATCCTCATTGCACCACCTTGACCTGGCTCAGCCTGTTT
CAAAATGGCTCAGTACAGAGCTGGGATTGTCCTGTTGCTTACAGAAGAGTACCTGCCAGAGAAAAGGCGGAATGAG
ATTCAACATGAACGGTCACTCCTATTCAATTGATGCTGGTACGAACGTGGGAGGTTAGGAGATGTAATGCAGTT
CGATTAAAGGTTCTAGAACAGGGGGTAGCAATGTCACGCAATTGGGGTCAAAATTGGCAGAGCAATGCATTACTGAT
GGTCAAATTCTATCATTAAAGTAACACAGGTGATGGCCGCACTGTTTGCAACAATGCTATCCCTGCTGGATGGTC
ATTGGGAAACTTACACAGGCGCACAGTTACTTGATTCCCTCAATTAGTATTGCAACTAAGACTATTACTACTAGTT
AATTAGTGTGGATGATGGGAAGAGCTTAAAGTGGCTCTAAACCTCTTTTTGTTGTTCTAAAGT
AAATTAAAGATATGAAGTTACGAGTTGTGTCCTGTGGTCAATTAGGAAAGACATAAATTAGGGGAGATTAGTTAGAGGAAAAGCTACTGTTCTGGT
GGATTATTGTACCAACCACCGCACATTACTCATTATTAGGGGATATTAGTTAGAGGAAAAGCTACTGTTCTGGT
TTGATGTTATTATTGTGGTGTGAAATTAAATTATGACTTGG

SEQ. ID. NO: 6

LeEXP10 (Expansin) predicted amino acid sequence. 250 amino acids

MSILWFSIIGVFCIQFINCVHGNEQGWIEAHATFYGGGDASGTMGGACGYGNLYSEGYGTNTAALSTALFNNGLSCGSCF
ELKCVGDSKWCLPGSIVVTATNFCPPNFALPNNAAGGWCNPPLHHFDLAQPVFQKMAQYRAGIVPVAYRRVPCQRKGGMRF
TMNGHSYFNMLVTNVGGSGDVNAVIKGSRTGWIAMSRNWQSNALLDGQILSFKVTTGDGRTVFCNNAI PAGWSF
GKTYTGAQFT

SEQ ID NO: 7

LeGAS2 (Gibberellin-stimulated/ABA down-regulated) cDNA from tomato seed, 489 bp

GGCACGAGGAGTTTTTTTTGGTATCTTAAACAATGGCATCCCTAAAGGGATTGCAGCTTGCTCATTGC
ATCACITGTGCTGTTCAATTCACTTATGCCCTCAAGAGGTAAATTAGCGGCAAACCACCGCACCAGTCCTAACCT
CCAAAGCCAATAGACTGTACGGGATCTTGTAAAACGAGGTGCAGCAAATCGTCACGGCAAATCTCTGCAACAGAGCAT
GTGGAAGTTGCTGCCGTACCTGCCACTGCGTGCACCAGGGACTCTGGAAATTACGAAGCATGCCCTGCTATTC
CCTTACTACTCATAACAGCACCCGCAAATGCTCTAAAAATTCAAGCATAATTGTGTTCTCGTCTGCAAATTAAAT
GTACTAGTATTCTTATTGAATACTACTATTATATTGTGTAATGAAACGTGAACGTGATTGAAATTAAAGTG
TAAAATTGAAGGTTG

SEQ ID NO: 8

LeGAS2 predicted amino acid sequence. 103 amino acids

MASLKGFAALLIASLVLVHFTYALQEVISGKPPAPSPQPPKPIDCTGSCKTRCSKSSRQNLCNRACGSCCRTCHCVPPGT
SGNYEACPACYFNLTTHNSTRKCP

SEQ ID NO: 9

LeARA (Arabinosidase) cDNA from tomato seed, endosperm-specific, 2113 bp

GGCACGAGGGAAAGATGGAGTCAAGGCATTCCATTCACTACGTACTGCTCTAGTTTGGTTGTCTGCCCTGTGC
 CAATGCTCTGCTACTGGGGTTGAAGCAAACCAGACAGCAGTACTGCTGTGAATGCATCCGAAGCATCAGCAAGGAGAA
 TACCTGATAACCTTTGGTATATTCTTGAGGAGATCAATCATGCTGGTGTGGATTGTGGCTGAGCTTGTCAA
 CAATAGAGGTTTGAAGGTGGAGGCCAACGTACCTCAAATATTGATCCTGGTATCATTGGAGATGAGTCCAAA
 GTGATTGTATCAACAGACCGTTCATGCTTGATCGGAATAAAATTGCAGTTCAAGTGCAGGTGCTGTGACCATA
 CAGGTGCCAATATCTGTCCAGATGGAGGAGTGGCATCTACAACCCGGATTCTGGGCATGAATAATTGAGCAGGGAAA
 GAGTTACAAACTAGTGCTTATGTCGTTCCGAAGAACATCAGTCAATGTATCTGTCGCTTAACTGGTTCAAATGGATTG
 CAAAAGCTGGCTGCCAACATTGTAGCTGCTGATGTTCAAGTTGGACGAAGGTGGAAATTGGTTAGAAGCAACAG
 GAACAGATCCAATTCAAGACTGGAATTGAGATCATCTAAGAAAGGTGTTGGATCAAGTCTCATTAATGCC
 TACTGACACATACAAGGGCATGGCTTCAGGAAAGATCTTTGGAATGCTTAAGGATTTAAAGCCAGCATTATTAGA
 TTTCCAGGTGGATTTGTTGAAGGTGACTGGTTAAGAAATGCATTGGTGGAAAGAAACTATTGGACAATGGGAAG
 AGAGGCCCGGACATTGGTGGATGTTGAAGGAAATTACTGGACTGATGGACTTGGACATTGGAGTTCTGCAGCTTGC
 CGAAGACTTAGATTCACTGCCGTGTGGTTCAACAATGGAGTCAGTCACCATGATCAAGTTGACACTTCCAGTATT
 TTACCTTTGTGCAAGAGATATTAGATGGCTTGAGTTGCAAGAGGTGATCCTACTTCACACATGGGTTCTATTGAG
 CCAAAATGGGACATCCAGAGCCTTGATTTGAGATATGTCGCTATGGAAATGAGGATTGTGGAAAAACACAATACCG
 TGAAATTACCTCAAGTTCTATACGGCCATCAAAGATAAATCCAGATATTAAAATACTCAAACATGTGATGGTTCT
 ACGAGACCACTGGATCACCCAGCTGATTATGATTTCATATTATAGCAGCGAAGTCTGTATTCTAATGCTC
 GCCATTGATAGTCACCAACGAGGACCAAGGCTTGAGTGGACTGCGTACTGGAAATGATGCTGGAAA
 AGTAGTCTTGTGGCATTGGTGAAGCTGGCTCTCATGGGGTAGAAAAGAACAGCGAAGCAATTGAAATGGCA
 AGTTACGCACCCCTATTGTTAATGACAATGACCGGAGGTGGAACCCAGATGCAATTGTCTCACCTCTCACAGATGT
 ATGGAACGCCTAGTTATGGATGCAGCACTCTCAAAGAGTCAAATGGCGTACTCTCTGAGTTGTCATTACAAGC
 TAATCCTCAAATTCACTTATAGCATCTGCCATCACTGGCGAAATTCACTGATAACAATGATTATTGAGAATAAG
 GTCGTGAACTTGGAACCACTGCCTTATAACTAAAATCTCTCACTGGATTGGGCAGAACCTCGTAGAGACGTTAT
 TTGGGGCTGTAATGACCGAGTTAACATCCAACATGTGATGGATGAAATTCTCTCAGAGAGCCTAATAAGGTAATACC
 TGTAAAACACAAGTTGAGAAAGTTAGTGACAACATGGATGTTGACTAGCTCCAAGATCTCTGAATTCAATTGACTTT
 TTATTAAGAAAATCAATAAACACAATGTTGATACTGCTCTGCTCTAAATCTCATGCTAAGTAGCTCAATGTATCA
 ATTAAATTAAAGAATATAATGTGGTTACTAATTATAATAATCATTTCATGGTTC

SEQ ID NO: 10

LeARA (Arabinosidase) predicted amino acid sequence. 674 amino acids

MESRHSIHVLLVLFGLSALCQCSATGVEANQTAVLLVNASEASARRIPDTLFGIFFEEINHAGAGGLWAEVLNNRGFE
 GGGPNVPSNIDPWSIIGDESKIVSTD RSSCFDRNKIAVQVQLCDHTGANICPDGGVG IYNPGFWGMNIEQGKSYKLVL
 YVRSEESNVSVALTGSNGLQKLAANIVAADVSSWTKVEILLEATGTDPSRLELRSSKKGVIWFDQVSLMPTDTYKGH
 GFRKDLFGMLKDLKPAFIRFPGGCFVEGDWLRNAFRWKETIGQWEERPGHFBDVWNWYTDDGLGHFEFLQLAEDLDSLPV
 WVFNNGVSHHDQVDTSSILPFVQEILDGLEFARGDPTSTWGSIRAKMGHPEPFDLRYVAIGNEDCGKTQYRGNYLKFYTA
 IKDKYPDIKIISNCDGSTRPLDHADLYDFHIYSSASSVFSNARHFD SAPRRGPKA FVSEYAVTGNDAKGSI LLAALGEA
 GFLIGVEKNSEAIEMASYAPLFVNNDRRWNPD AIVFTSSQMYGTPSYWMQHFFKESNGATLSSSLQANPSNSLIASAI
 TWRNSLDNNNDYLRIKVNVFGTTAVITKISLTGLGQNSLET LFGAVMTELTSNNVM DENSFREP NKVI PVKTQVEKVDNM
 DVVLAPRSLNSIDFLLRKSINNNVDTASVLKSSC

SEQ ID NO: 11

LeMAN2 (endo- β -mannanase) cDNA from tomato seed, 1480 bp

GGCACGAGAAATAATAATGGCTTATTTCAAAGACTAATTAGTTGTATTTGTGCTTTCTTTGTCCTTAGCTTTG
 CATGTGAAGCTAGGGTTTACTTGATGAAAATAATGCAAATGATCAAGGGTTGTTAGAGTCATGGTGCACATTTGAA
 CTCAATGGATCACCTTCTATTCAATGGTTCAACTCTTACTGGCTAATGCATGTTGCTGAGCCTAGTGAGAGGTA
 CAAGGTCTCCGAGGTCCCTCGCGAACATCTCTGCAAGGGTTCTGTATGCCGTACTGGGCTTTAGTGATGGAGGTG
 ATAGGGCATTACAAATTCACCTGGTGTATGATGAACGTGTTTCAGGGTTGGATTTGTTATATCGGAGGCTAAG
 AAGTATGGAATTGATTAATCTTGAGCTTGAAACAACTACAACGATTTGGAGGGAAAGCTCAATATGTTCAATGGGC
 GCGAAATGCAGGAGCTCAGATAATGGAGACGATGATTCTACCACTTAATTATATTACCCAAAAATTATTACAAGAACATC
 ACATTAAGAAAGTTGTTACAAGGTTCAATACCATTACTGGAATGACATATAAAGATGATTCAACTATCATGGCATGGAA
 CTCATGAATGAGCCTAGGAACCAAGCTGATTATTAGGAAATACTCTAACGCATGGGTTCAAGAAAAGGCAAGTTTGT
 GAAATCACTTGATAACAAACATTGCTGGAAATAGGAATGGAAGGATTTATGGTGATTGAGCTGCCAGAAAGGAAGTC
 TTAATCCTGGCTATCAAGTTGGAACAGATTTATTAGTAACCATTGATCAAAGAGATTGATTGCACTATTGCA
 TACACTGACCAATGGTTATCTGGTCAAAGTGATGATGCTCAAATGATCTCATGCAAAATGGATGACAAGTCATTGGCA
 AGACGAAAAACATACTTAAAAACACTGGTTCTAGCTGAATTGGCAAATCAAGTAGAGATCCAGGATATAATCAA
 ATATACGCGATACATTATGAGCACGATATAGAAATATTATAGTTAGCAAAAGATGGAGGAACGATGGGAGGAAGT
 TTAATATGGCAGCTCGTTGCACAAGGCATGGAAAATTATGAAGATGGTTATTGATTGGAAATTGGGAAAAATCCATCCMC
 TGCTGGAATTATTACAAGTCAGTCTCATGCCATGACAGCTTGGCTATTGGTAAGATTAGGACTAATTGATAAGT
 AATACATTAGAGGACATGGAAGAAAATAAGTATACTTTAGTTAGTCCATCCCTTTTTGGTAAATTATGAA
 TATGTTGTAATAATTGGCTAAGTATATGTAACCTTCAAAATAATCAATTGAAATGTATGTTAACTCCTTATATAT
 ATGTAAGGAA

SEQ ID NO: 12

LeMAN2 (endo- β -mannanase) predicted amino acid sequence. 415 amino acids.

MAYFQRLISCI FVLFLSLAFACEARVLLDENNANDQGFVRVNGAHFELNGSPFLFNGFNSY WLMHVA AEPSERYKVSEV
 LREASSAGLSVCRTWAFSDGGDRALQI SPGVYDERVFQGLDFVISEAKKYGIRLILSFVN NYNDGGKAQYVQWARNAGA
 QINGDDDFYHLIILPKNYYKNHIKKVVTRFNTITGMYKDDSTIMAWELMNEPRNQADYS GNTLNAWVQEMASFVKSLDN
 KHLLEIGMEGFYGDSPVERKSINPGYQVGTDFISNHLIKEIDFATIHAYTDQWL SGQSDDAQMI FMQKWMTSHWQDAKNI
 LKKPLVLAEGFKSSRDPGYQNQNI RDTFMSTIYRNIIYSLAKDGGTMGGSLIWQLVAQGMENYEDGYCIELGKNPSXAGIIT
 SQSHAMTALAHLVKI

SEQ ID NO: 13

LeXET4 (Xyloglucan endotransglycosylase) cDNA from tomato seed, 1045 bp

AGAACAAAGGAAGATAGTAAGTAATGAATATGAAGGGAGTTTAGTGCTTTGATTAAATTGTCAATAATTGGCA
AGTTGTGGGGCTCCAAGGAAGGTAATTGATGTGCCTTTGGAACAACATGAACCAAGTGGCTAGTCACCATAATTAA
GTACCTTAATGGTGGTACTACGGCTGAACCTCTTCTGACAAATCCTCTGGAACATTCAATCAAAGAGATCGTATC
TATTGGTCATTCAAGCATGAAAATGAAGCTTGTGGAGGAGATTCTGCTGGTGTCACTGCTTTATTATCATCG
ACTAATGCTGAACACGATGAGATAGATTCAATTCCCTCGGAATAGAACCGGTAGCCATACATATTGCAGACTAATGT
GTTCACAGGAGGCAAAGGAGACAGAGAACAGAGGATCTATCTTGGTTGATCCAACCAAGGACTTCATTCAATTCTG
TTCTTGGAACACTTACCAAATTGCGATTGGATGATGTCCAAATAAGAGTATTCAAGAATTCAAAGACATAGGA
GTGAAATTCCGTCAATCAGCCAATGAAGATCTACTCAAGCCTATGGAACCGGGATGATTGGCTACAAGAGGAGGGTT
AGAGAAAACAAATTGGTCTGGGGGCCATTCACTCGCTTCTATACTTCAATTCCACATTGATGGATGTGAGGCTGTACAC
CACAAAGAGGTACAAGTTGTAACACCAATGGCATGAAATGGTGGGATCAAAGGCTTCCAAGATTAGATGGCCCTGAA
TATAGAAAACCTCATAGGGTTAGACAAAATTNCATAATATAACTATTGTAAGTACAAAAAGGTACCCCTACACTTCC
TCTAGAGTGTACAAGGGATAGAGATCTTAATTAGTTCTTATATGGGGCATTTCATTACTCTTATATAAAA
ACTAATCCAGTTGTATGGAGACGAATTCAAATTATAAAATTGGACTGTATTGAAGTTCTGAATATGTCAAATTAAAG
ATTTT

SEQ ID NO: 14

LeXET4 (Xyloglucan endotransglycosylase) predicted amino acid sequence. 295 amino acids

MNMKGVLVAFVLINLSILASCAPRKVIDVPFWNNYEPSWSSHIIKYLNNGTTAELLLDKSSGTGFQSKRSYLFHFSMK
MKLVGGDSAGVVTAFYLSSTNAEHDEIDFEFLGNRTGQPYILQTNVFTGGKGDREQRIYLWFDPTKDFHSYSLWNTYQI
AI FVDDVPIRVFKNSKDIGVKFNFNQPMKIYSSLWNADDWATRGGLEKTNWSGAPFIASYTSFHIDGCEAVTPQEVCN
TNGMKWWWDQKAQDLDGPEYRKLHRVRQNFXIYNYCTDRKRYPTLPLECTRDRDL

SEQ ID NO: 15

LVA-P1 (vacuolar H⁺-ATPase) cDNA from tomato seeds. 900 bp.

TTCCCCCTGAACCCCCCGGTGAGGGAGGAGAAAGAGATCTATTCAACTCCAAAACCTCAA
AATTCTCAGATCCAAATCCTAACGAAAATGTCGAACCTTGCCGGAGATGAAACTGCTCCC
TTCTTCGGCTTCCCTGGCGCCGCGCTGCCCTCGTCTCATGTATGGGGCAGCTTAT
GGAACAGCAAAGAGTGGTGTGGAGTGGCGTCAATGGGAGTGAATGAGGCCAGAGTTGGTG
ATGAAGTCCATTGTGCCAGTGGTTATGGCTGGTGTAGGTATTTATGGCTGATTATT
GCTGTGATCATCAGTACTGGGATTAACCCAAAACAAAGTCGTATTACCTATTGATGGC
TATGCTCATCTCATCTGGTCTGCTGGTCTGGCTTTCTGCTGGAATGGCT
ATTGGTATTGTTGGAGATGCTGGTGTAGGGCTAATGCACAACAACCAAGCTTTTGTC
GGAATGATCCTCATTCTCATTTCGCTGAAGCCTGGCTTTATGGCTATTGTTGGC
ATTATCTTGTCTTCCCGAGCTGGCAGTCTAGAGCCGAGTGAAGTTAATCCATTCTTAC
CGCACTGTGTGGTTCTGAAGACCAAGACAGCTAAAGCCTAAAGTCAGAGATCTAATA
TGTGTATTGTTATTGACACCACAGCTGCCACTTTCTGTGTATGATCTGTTGAGA
GTAGGAATTCTCTTTCTACTTAATAATAGCTAAAGAGCTGTGCAATTGGTCTGTAT
TTTGTGTATTGCACTCATTATTGAGACAGTTGAGAACTATTTATTTCTAAGATTG
TGCACGTATGAACCACTCTCATCTATATACCACCATGGTTATTCTAAAAAAAAAA

SEQ ID NO: 16

LVA-P1 (vacuolar H⁺-ATPase) predicted amino acid sequence. 164 amino acids

MSNFAGDETAPFFGFLGAAAALVFSCMGAAYGTAKGSGVGVASMGVMRPELVMKSIVPVVM
AGVLGIYGLIIAVIISTGINPKTKSYLFDGYAHLSAGLACGLAGLSAGMAIGIVGDAGV
RANAQQPKLFVGMIILILIFAEALALYGLIVGIIILSSRAGQSRAE

SEQ ID NO. 17

Cel55 (cellulase or β -1,4-glucanase) cDNA from tomato seed, 2003 bp

TGGCTACTATGACCGGGGGATAATGTGAAATTGGTCTACCAATGGCATTCACAGTTACTATGATGTCATGGAGCATA
ATTGAGTATGGGAAGCAAATGAGTGAAGTGGAGAGCTTAGTAATGCTATTGATGCTGTTAAGTGGGGTACTGATTATC
TACTTAAAGCTCATCCTGAACCACATGCTCTATATGGAGAGGGTGGAGATGGTACCCACAGATCATTACTGTTGGCAAAG
ACCAGAGGATATGACCACTCAAGAGCTGCTTACAGGATCGATCCAAGTGGCACGAGGTCGGATCTGCCGGAGAAACA
GCAGCCGCATGGCGCTGCTTCCATCGTCTCGTTACAACCCGGATATTCTAATGAGCTCCTTAATCATGAC
ATCAGCTGTTGAGTTGCTGATAAGTATAGGGCAAATATGATAGCAGTATTACTGTCGCCAGAAGTACTACCGATC
TGTCACTGGATACCGGGATGAATTACTATGGGAGCTGCATGGTATACAAGGCATCAAACAATCAATTTCATTGAAT
TATTGGGAAGAAATGGTATGCTCTGGGAACTGGTGGCTATGACTGAATTGGTGGGATGTCAGTATGCTG
GTGTCCAACACTGTTGCTCAGTCTTGATGTCGGTAAGGCTGGCATAATGCACCTGTATTGAGAAGTACCA
GAAAGCAGAGAACCTTATGTGTTCAATGCTGGAAAGGGTAATAGAAATACCCAGAAAACCTGGAGGTCTCATCTAT
AGGCAAAGGATGGAACAATATGCACTTTGTTACTAGTGCTGCATTCTGCCACTACTTATTCTGACTATTGGCTTCT
GCTGGCAAATATCTCAAGTGTCTTCCGGTTCTCGGAATGAGCTCCTCTCATTGCTAAGTCACAGGTGGACT
ACATTCTGGTACAATCCGAGAGCAACAAGTTACATGGGGATAATGGAAACAATTACCGAGACAAGTACACCCAG
AGCTTCTCCATTGTCGTTAAGGTTAACCTAGTTGTTAGCTGTCGGGAGGCTATGCTACCTGGTATAGT
AAGGCGAGCGATCTAACATCTCTTACTGGTGCCTTGTGGACCTGATGCCTATGACAACCTTGCTGATCAAAGAG
ATAACTATGAGCAAACGTGAGCCTGCCACTTACAACAATGCTCCGGTATCGGAGTGTAGCAAGACTTCATGCTGGTCA
CAGCGGTTACAATCAGCTCCTCAGTTGTTCTGACCCGAAGCCAACGCCAAAGCCAGCTCCAAGAACTAAAGTAAC
CCAGCTCCAAGGCCAACAGACTTCCAGTCCCAGCTAACGCTCATGTTACTATTCAACAAAGGGCAACTAGTTCATGG
CTCTGAATGGGAAGACTTACTACAGATACTCAGCAGTTGTAACCAACAAGTCCGGAAAGACGGTTAGAAACTTGAAACT
CTCCATAGCTAACGCTCATGGCCTCTGGGGCTAACAAAGTACGGTAACCTGTTCATCTCCAGCTTGGCTAAC
TCTTACCTAGCTGGTAAAGCCTAGAGTTGTACATTCACACTGCTCACCTGCAATGTCCTCCGGTTCAAGCTACA
CTCTAGTCTAAAGACATCAACAAGTTGAAAAAGAAAAATGAAGTGGAAATGAAATGTCACACATTTCCTTGTGG
TTCTAATGTAAGAATTGTCAGCTACTACTACTACTAGTACTAACGCTAACGTCTGGTTGATTGTTAAGACTTTGATAAC
GTTGATAATGTGGAGAAGTCGGAGGAAGAAGGAGTGAAGAGACGAGTTATACAAGTGCTCATCCTCTTCTCC
CTTATATATACACATAGTGTTCAG

SEQ ID NO. 18

Cel55 (cellulase or β -1,4-glucanase) amino acid sequence. 253 amino acids

MSWSII EYGKQMSSEGELSNIDAVKWGTDYLLKAHPEPHVLYGEVDGTTDHYCWQRPEDMTTSRAAYRIDPSGTRSDL
AGETAAAMAAASIVFRRYNPGYSNELLNHAHQLFEFADKYRGKYDSSITVAQKYYRSVSGYADELLWGAALYKASNNQF
YLNYLGRNGDALGGTGWSMTEFGWDVKYAGVQTLVAQFLMSGKAGHNAPVFEKYQQKAENFMCMLGKGNRNTQKTPGGL
IYRQRMEQYAVCY

SEQ ID NO: 19

Cel68 (cellulase or β -1,4-glucanase) cDNA from tomato seed, 1988 bp

GGCACGAGCCCTGTTCTATCCCTCAACTACCTCAAATCTCTTCTTTAATTAAACCATTTCATTTCACCCCTAAA
 TTTCTCCATTATTTCTCATTCTCCAATCTTTCTTGTGTTCTGTTATATTATAATTTCACCTGCA
 TTGCATATTAAAAGAACAAAGTTTCAGCCATATAATCCCGATTGTGTGCGGGATCATTAAATCAGGTAAATTG
 TGAGGTGAATAAGGATGGGAGAGAAATCAAAAAAGGGGTTGGTGTGGATGGATTATAGCTTGGTAGTGGTTGCTGC
 GGCTGCTGGAGCCATTGTGGTACTTTATGAAGAACATAATGGTCTGGCTCAGATGCTGCTCTGGTCTGAA
 AAGAAATATGGAGATGCTTGAAAATTGCAATGCAGTTCTTGATGTCCAAAATCTGGGAAATTGGTAAATAATAAGA
 TATCATGGAGAGGCATTAGCCGTCAAAGATGGAAGTCAAGAAAAGTTAGACCTCAGTCATGGATACTACGATGCTGG
 CGACCACATGAAATTAAATTCCCAATGGCTTACACTGCCACTGTGTTGTCATGGGCATCCTGAGTATGGTAATCAG
 ATGAAGGGTGTGGGCAGTTGGAACCTGCTGAAGACTCACTCAAATGGATTACGGACTATCTTATCAATTGTCATCCAA
 AAGATGATGTTCTTATATTCAAGGTGGGTGATGCTGATGCTGACCATAATGTTGGACAGACCCGAGGACATGACTGA
 GGCAAGGCCTCTAATTCAAGATAATGCTTACCCCTGGAACAGAACAGTTGCAAGCTGAAACTGCTGAGCTATGGCAGCA
 GCATCCCTGGTCTCAAGTCGAAGAACTCAGCATACTCAAGTAACCTGCTTAAGCATGCTAAACAGTTGTTCACTTTG
 CTGACAAACATAGAGGTACCTACAGCAGAAATATTCCCTGAAGTCGCAACATATTACAATTCAACTGGATATGGAGATGA
 GCTCTTGTGGCAGCAAGTTGGCTATATCATGCAACTGGGATCAAATATATTGATTACGCACTGGAAATGCA
 GATTCTTTGGTAATTGGAAATCAACTTGGTTAGCTGGATAACAAACTAGCTGGTACTCAGGTCTTTGTCCC
 GGGTCAGCTTCTTAATTGAAAGTCTCAAACACTCAGACACACTCAACAGTACAAAAAAACTGCAGAAGCTGTAATGTG
 TGGTCTCTTACAAAATCTCAAACAGCTACATCCAGCAGAACTGATAGTGGTTGATATGGATAACTGAATGGAACGCG
 TTGCAAGCATCCTGTAAGCCTCTGCCTCTTAGCTGTGTCAGTGATTATATGCTCACTCCAAGACTGATAAAATGA
 CTTGTGATGGAATGAGTATACACCATCTGATCTCAGAAAGTTGCCATGTCCAGGCCATTATGTATTGGGTGACAA
 TCCAGCAAAGATGAGTTATCTCGTAGGCTATGGGACAAATATCCGCACTGTTCAACCACAGAGGGCTTCCATTCT
 ACAGACGCCAATACTAATTGCAAAGAAGGTGGAAGTATCTAGACTCAACTGAACCAAATCTAACGTTGCAACTGGGG
 CTCTCGTTGGTGGTCCATTCTTAATGAGACATATATCGATTCAAGGAACAACCTGGTTCAAGGAGAGCCAACCACATA
 CAATAGTGTGTGGTTGTTGCCCTTCTTCTGGTTGTTCCACTTCCTCAGTGGTTCAATCTTACCTGAGCTGGA
 TACCATCTCACGTCGTACGTCTTATCATACTGATGCATACATAATTCTTATTGTTAGTGTGTGTCACATCAACTGTTG
 TAAAGGTTGCTTATATATATGTGTATGTTAGTGTGTACTTATATCAAATATA
 TGTGCCGAACAGG

SEQ ID NO: 20

Cel68 (cellulase or β -1,4-glucanase) amino acid sequence. 519 amino acids

MGEKSKKGGCGWIIALVVAAAAGAIIVVLFMKKHNGSDAASGASEKKYDALKIAMQFFDVQKSGKLVNNKISWRGD
 SAVKDGSQEKLDSLHGYYDAGDHMKFNFPMAVTATVLSWAILEYGNQMKGVGQLEPAEDSLKWITDYLINCHPKDDVLYI
 QVGADADHKCWDRPEDMTEARPLIQINASTPGTEVAAETAAMAAASLVFKSKNSAYSSNLLKHAKQLFTFADKHRGTY
 SENIPEVATYYNSTGYGDELLWAASWLYHATGDQIYFDYATGKNADSFGNFGNPTWFSDNKLAGTQVLLSRVSFFNSKV
 SNSDTLQQYKKTAEAVMCGLLPKSPTATSSRTDSGLIWITEWNALQHPVASAFLAVLYSDYMLTSKTDKMTCDGNEYTPS
 DLRKFAMSQANYVLGDNPAKMSYLVGYGDKYPQYVHHRGASIPTDANTNCLEGWKYLDSTEPNPNVATGALVGGPFLNET
 YIDSRRNNSVQGEPTTYNSAVVALLSGLVSTSSVVQSFT.

SEQ ID NO: 21

LeXPG1 (polygalacturonase) cDNA from tomato seeds. 1603 bp

TAGCCGACAGTGATTTGACCTTGTCTAGTACTACTCTGTTAAGTTACAAAAGTGCTTCCCACT
AATGGAAAATGGGAAAGATGGCATCCTCCTCCATTCTTCCTTAATTCTCGTCCTATTCCAA
ATCTGTTGCTCTGTTCTGTCACTCTACAGTCTCTATTCCCTTGAGTTCTGCAAATGTAAGT
GGAGTTGAGTCACTCTTACAGCTACTCCAGCATCTGCTTCTTGAGGACCAAATCGGAATCCC
TTTTCGAGTCAACGATTGGAGCTACCGGAGATGGGATTACCGATGACACTAAGTCTTTAA
AGATGTTGGGATATGGCCTGCTCGTACCATCACATGCAAAGATTGTTATCCCTGCTGGTTAT
TCTTCTTAGTCCGACAAATTAAATTGCTGGCCTGCCGGTCAAAGGTGTCTATACGGATTG
CAGGTACTATTTAGCACCAAAGGATCCTGACGTCTGGGATGACTTGAATCCACGAAAATGGAT
CTATTCTTAAAGTAAAACACCTGACAGTAGAAGGAGGAGGAATTATAGATGGTATGGGCCAG
GAGTGGTGGGCTCGGTATGCAAGGTCAACAGAACAAATCCTGTCATCATGCTCCAACGGCTT
TAACTTCCACAAATGCAACAAACCTGAAGGTCAAGAACATAAAAGATCTTAAATAGTCAACAAAT
GCATTTAGCATTACTGGTGCAAACATGTTACAATATCACAACCTCGTAGTCAAAGCCCCAGGT
GATAGCCCTAACACCGATGCAATCCACATAAGTTCATCACAGGTCAATGTCAAGGATTGCA
TTATTGGCACAGGAGATGACTGCATATCTATTGTCGGCAATTCATCACGGATCAAAGTCAAAGA
CATTGTGTGGGCCAGGCCATGGTATAAGCATTGGAAGCTGGAAAGTCAAATTCTTCT
CAGGTTACAATGTTATGTTAATGGAGCATCTATTCCAACACTGAGAACGGGTTAGAATAA
AGACCTGGCAGGGTGGTCTGGGTTGTCAAAAAGGTTCTTCGAGAATGTTGGATGGAAAA
TGTCTCAAATCCTATCATAATCGACCAATATTATTGTGATTCTAGGAAACCTTGTCAAACAAAG
ACTTCAAACATTCACATTGACAACATATCCTTATGGTATTAAAGGAACCTCAGCTACAGAAA
GGGCAATAACACTAGCCTGTAGCGACAGCTTCCCCTGCAGAAGGTTGTACTTGGAAAGATATTCA
ACTTACTTCATCTCTGGGATCCTACAACATTCTTGTGGCAGGCATATGGCACAACCTCA
GGTTAAATTATCCCCCTCTGCTTCTGTAATGACGGCATTCTCAGCCAAATTCTTAT
CTAACTGGAGTCAATCGATATGATGTTCTATTGTAAGTGTACAAATGTCTCCACCATAATC
ATCGAAAGCTGCTAAATTGTTGTCATTGATATTCTATTCTGAGCGAAGCTAGTTAACTTAT
CCAGAGGAACAGTATAGCCATGACGTACTGATGAGTAGCTTGTCCAAGTACTAATGTTATGG
AAG

SEQ ID NO: 22

LeXPG1 (polygalacturonase) amino acid sequence. 452 amino acids

MGKMASSSISSFNFSSYSKSVALFLSLYSLYFLLISANVSGFESLLQLLPASASLRKSESLFRVNDFGATGDGIID
DTKSFKDWDAMCSPSHAKIVIPAGYSFLVRQINFAGPCRSKVSIRIAGTILAPKDPDVWDDLNPRKWIYFFKVKH
LTVEGGGIIDGMGQEWWARSCKVNRTNPCHAPTALTFHKCNNLKVKNIKIFNSQQMHLAFTGCKHVTISQLVVKAP
GDSPNTDAIHISSTQNVKDCIIGTGDDCISIVGNSSRIKVKDIVCGPGHGISISLGKSNSFSQVYNVHNGASI
SNTENGVRIKTWQGGSGFVKVVFENVWMENVSNPIIIDQYYCDSRKPCSNKTSNIHDNISFMGIKGTSA
TERAITLACSDSFPCRRLYLEDIQLTSSSGDPTTFFCWQAYGTTSGLNYPCCFPNDGILQPKFLSNWSQSI

LeXPG1 promoter sequence. Total 4267 bases of BamHI digested fragment from genomic library.

GGATCCAGAGGGCATAGTCTATTAGGCTCTGGACGCTGGTAATCCTATGTG TACGAACGTATT
 TTATTCAAAAAAAAATCTTTATTCCATCTTGAGCTGTTAGAATCAATTCAATGTGCG ATGATTCAATCTGA
 AATTAAGAATATTTGCTTATCTTATTATAAGTGATACGTTAATTGTTCCCCAACTCAAT TCGCACACTTTT
 TTTTAAGTATATTTGTTGATTTCATTTAATTGACTTGTGAAAATTGAGTTAGTAGGCCGAAGAGAGAATAAT
 TGTCAAAAATAATATAAGTAGATCAAGAACCTTAAATTATTGATGGAACCTCAACACACCAACATTAAAACCTGAA
 TAATATAAGAGGGACCCAACATCTAGAACCTATGAGTGAGAAAAATTGATCAAATCATAAAAGGTGATCAAGGACC
 TTTGGACGTGAGACTACAACAAGTTGAGTATATTCCCTCTGAGGCTGGCATTGACAGGATCAATCTCAAGAAGTGAAG
 TTAACATAAAAAAGAAGAAGTAGAAGATTAGTAAGAAAGGGAAACTTATCAATATAGCATAACAATCACTGCT
 CTATTGCTCTGAGCCCTGCTCGATTTTTAATTGTTGAGTCGATGTCTATTAGACAAACCTCTAACTC
 ATAAAGTAGTGGTAAGGTCTATACATCCTATTCAAATTGTTATGCGATTACACTAGGTATACTGTGGTGTGA
 GGTAAAATTAAATAGAGAATTTTATGGCCCACATTGTTGAAAAGACAGGGAGACCAAAGAGTGTGCTGCAATTATG
 AAAGATTAACATTCTGTACCCAACTGTGTCTATCACAAGCTACAAAACAAATTGCTGCCCA TGCACGCAGCCCTC
 TCGTTCCCATATATTAAATATCTATGTTCTAGAAGAGCCACTGGTGTGAGAACATTGCTAGCAAAGAAACATCCT
 GTTCGCTAAGTAGAGAGCATGTTCAATCTCAAGCATGGGAATGCTTACCTCTCAATAGTACTTTTTGCCTTGT
 TATAAAACATACGAACATGGAATGGAGCTAGTAATTGACAATTGTAACATGGAAGTGTGAATATCTC
 AACATCACATATTCCCGAGATCTAACATGTACGACTGACGGCTGACATGCTATATTGAGGCCGAAACACATTCC
 AGATCATGCAATTGCTATTGTTCCCAATTGAAAGACCTAACTGATCTCATGTCAGATCGTTATTCTGTGCTCCT
 CATTGTTGACATTCAAATCTGGTTAATTGCTACTCAGTACCGTTAACTGGACATCTTAAATTACCACTG
 TTGCAATTGAACTGTGTGATTGAAATGTTGCTGCTCTCAAGTGCAGGTGTGACAATTCTCACTTATTGTTGATT
 GGCAAGCATGTGAAAATTCTTTGATTGGATGACAGTGAGATTCAACTCAAGACTTTGCTGCTGCTGATACCAT
 GTTAGAAGTGATGATCATCTTATCTAAAGCTTAAAGCTTAAAGAGAGTACATTGTTATGTTGATGTTGATT
 ACAAGTAATAATGAGGAAGCACACTCACAAAGTAAGGTGTACAAATTACAAATTGCAATCTCAATTGTTCTTAAA
 GCGAGCTCTTGAATAACCTTATTGTTAAAGAAAGCGTAAGACATCAGGTATAATTGAAATCATTGTTCTT
 TGTTTCATTGCCAATTATGCGAAATGGCTCATGCTCAATTATAATTAAAGGATTATAACTTCTATCACTAGATT
 TGATTAACCTCTTTTATCCTTATTCTACATGACTCAAGTCTCAACACATTGTTAACTCAAAACTCAATAAGTTT
 GTATTCTTTATGTTAGAAAACATATTGTTATTGTTACTATTAAATTATACCATCAAACAGTAAGTAATGACT
 TTAACATATATCATATATAATTAAATAGTGAATAATTACAAATTGGTGAATTGATGTTCTAATAGAGAGA
 TTAAAAAATGAAAATTAAATTAAATAAAAGCCGTTGATAGTTAATTGAAATTGATGCAATTAAATTGTTATGAA
 GTAATGCCACATTGGTATTGATTGAACTTAGGAGTTACTCAAGCATAAAATTCTACAAATGTTAATTGTT
 TATTAGAAAATTCAACATCTAATATGTGAATGTTAAGTGGTCAGAATTGACTAGAAAATCAAAAGTCTATAAT
 TTTTTTATTAAACTATTGTTAGTTAAAGTATTAAATTAAAGATAAAATTGTTAAATTAAATTAAATT
 TACATAGGTAAAATTAAATCTAACCGAAAGAATTGTTCTATTGATAATTGCTAATTACTAATTCTCATAACTAATAAGT
 GCTTGTATAATTCAATTATGAAACCGCTTTAAATCTAAATTGTTACTATCTACTAAAGGCCGGTAAAAAATT
 AAGTATCGATTCCCACCTGGAAAAATCACATGATTACGCTTTCGAAATCATTGTTGAGTCAGTTAGATGATGT
 TTATTGTTAAGTGTAAATTGAAAGGCTAAACTTCAAAATTAAAGTCAAAGTACATCTATGTTGATGTTCTCTCA
 TTAAATAGGCTGTGACTTATGAAGCCTTTTTGGGAAAATGTAGCCGACAGTGATTGACCTTGTCTAGTACT
 ACTCTGTTAAGTACAAAGTGTCTCCACTAATGGAAAATGGGGAGATGGCATCCTCCATTCTCTTAAATT
 TCTCGTCTATTCAAATCTGTTGCTCTGTTCTGTCACTCTACAGTCTCTATTCTCTTGTGATTCTGCAAATGTAAG
 TGGATTGAGTCACTCTACAGCTACTCCAGCATCTGCTCTTGTGAGGACCAATCGGAATCCCTTTGAGTCAC
 GATTTGGAGCTACCGGAGATGGGATTACCGATGACACTAAGGTGAATCCCTTATTCCGCTAAAGGTCTTAAATT
 ATCGGCATTCACTATATGATTATGTGCTTCATTAAAGAGCCGCTGAAAAATTGTTATTGTTGATGTTCTGGA
 TCATTCTCGATCTATCGAAATAGCTCTCCACATATACATTACCCCTCTATACCCCATCTATACCCACTATGAGAT
 TATGCTGAGTATGTTGATTGTTGCTGATATCGATCAAAATTGTTATTGTTGATGTTCTGAGTATT
 GGTGTTACAAGACAACGACTAATGAAAGTAATAAACGAAAAGAAGGAACACAATGTGAAACCTTCAAGGGAGAGA
 ACCAATGCCAATCACCAACTCTTATTATTCTCAATGAATGATAGATAACAGATGCACCAATAACCTGTTGATA
 GTTGGTATAAGAAACTATCTAGAATGTTAAGGATGTCAACCCCTGAGCCAGGCCACAAACGTAACAATAAGTT
 TGTTATTGAAAGGATGATTACTCATTCTTTGTTACTGTAGTAAGATTGATACATTAAATTAAAAAAATAATTAGA
 AAAGGTTTGAAGAAGTATAGAGCATATATTGTTATTGTAATGGGAAACTAAGCAAAACTGCTTGATTGCTGAC
 TTAGTTCAAGAACAAACATAACTGTGAAATCCACTAAGTGGGTTGGGGAGGGTAGAGTGATACATAGCTTAC
 CATTACCTTGTGGAGGTAGTGAATAATTGTCCTTATTAAATTGTTGAGCTAAAGTTTTGAGACTTTTTATGAA
 GATGGATCACCAGAAATGAGAAATACTTAGCTTCAGTTAAGAAGTTTTTTGCTTCACTTAAAGAT
 GTTGGGATATGCCCTGCTCGTCAACATCACATGCAAAGATTGTTATCCCTGCTGGTTATTCTTCTAGTCCGACAA
 TTAATTGCTGGCCTTGTGGTCAAAGGTGTCTATACGGGTGAGTGCTTAATTGAGATATTGCACTTGTCTAAC
 TTTATTCCACGAAATGGTATTCAAGGTGTCTTTGATATACACAAGATTGCAAGGTACTATTAGCACCAAGGATCC
 A

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US00/26884

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : C12N 5/04, 5/10, 15/00, 15/09, 15/63, 15/70, 15/74, 15/82, 15/87; C07H 21/04; A01H 1/00, 5/00, 9/00,
11/00
US CL : 435/ 320.1, 419; 536/ 23.2, 23.6; 800/ 290, 295, 298, 317.4

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 435/ 320.1, 419; 536/ 23.2, 23.6; 800/ 290, 295, 298, 317.4

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Please See Continuation Sheet

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	SHCHERBAN et al. Molecular cloning and sequence analysis of expansins - a highly conserved multigene family of proteins that mediate cell wall expansion in plants. Proc. Natl. Acad. Sci. USA. September 1995, Vol. 92, pages 9245-9249, especially page 9245 column 2 4th full paragraph, page 9246 figure 1, page 9247 column 2 paragraphs 2 and 4, page 9248 figure 3.	1-14
Y	COSGROVE, D.J. Plant cell enlargement and the action of expansins. Bioessays. 1996, Vol. 18, No. 7, pages 533- 540, the entire article.	1-14
Y	ROSE et al. Expression of a divergent expansin gene is fruit-specific and ripening-regulated. Proc. Natl. Acad. Sci. USA. May 1997, Vol. 94, pages 5955-5960, especially page 5956 column 5th full paragraph, page 5957 figures 1 and 2.	1-6
Y	SHIEH et al. Expansins. J. Plant Res. 1998, Vol. 111, pages 149-157, the entire article.	1-14
Y,P	BRADFORD et al. 'Gene expression prior to radicle emergence in imbibed tomato seeds.' In: Seed Biology Advances and Applications, Proceedings of the Sixth International Workshop on Seeds. Edited by Black et al. Merida, Mexico, 1999, pages 231-251, especially pages 238-240.	1-14
T,E	CHEN et al. Expression of an expansin is associated with endosperm weakening during tomato seed germination. Plant Physiology. November 2000, Vol. 124, pages 1265-1274.	1-14

Further documents are listed in the continuation of Box C.

See patent family annex.

Special categories of cited documents:	
"A" document defining the general state of the art which is not considered to be of particular relevance	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"E" earlier application or patent published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

26 December 2000 (26.12.2000)

Date of mailing of the international search report

25 JAN 2001

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703)305-3230

Authorized officer

Cynthia Collins

Telephone No. (703) 605-1210

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US00/26884

C (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	WO 00/12715 A1 (ADVANCED TECHNOLOGIES LIMITED) 09 March 2000, page 17 2nd paragraph, page 18 2nd full paragraph, page 20 2nd full paragraph through page 22, page 23 paragraphs 1-2, Figures 1 and 2, sequences listing pages 1-8.	1-10
X	US 5,929,303 A (BENNETT et al.) 27 July 1999, column 9 lines 21-67, columns 10-11, column 12 lines 1-6, column 13 lines 1-67, column 14 lines 1-16, sequence listing columns 17-28.	1-11
X	US 5,952,543 A (FIROOZABADY et al.) 14 September 1999, column 7 lines 22-43.	9-10
Y	BRUMMELL et al. Differential expression of expansin gene family members during growth and ripening of tomato fruit. Plant Molecular Biology. 1999, Vol. 39, pages 161-169, especially page 164 Figure 1, page 165 Figure 2.	1-14
X	BRUMMELL et al. GenEmbl accession number AP059488	1-4
X	BRUMMELL et al. SPTREMBL accession number Q9ZP32	1-4
X	D'ASCENZO et al. GenBank accession number AI781569	1-4
X	LINK et al. SPTREMBL accession number Q9ZP37.	1-4
X	LINK et al. GenEmbl accession number AP049353	1-4
X,P	US 5959082 A (COSGROVE et al.) 28 September 1999 columns 29-30, and attached SEQ ID NO:2.	1-4
Y	LINK et al. Acid-growth response and α -expansins in suspension cultures of bright yellow 2 tobacco. Plant Physiology. 1998, Vol. 118, pages 907-916, especially page 912 Figure 6, page 913 Figure 7.	1-14

Form PCT/ISA/210 (continuation of second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US00/26884

Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. Claim Nos.: because they relate to subject matter not required to be searched by this Authority, namely:

2. Claim Nos.: because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. Claim Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)This International Searching Authority found multiple inventions in this international application, as follows:
Please See Continuation Sheet

1. As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1-14

Remark on Protest

The additional search fees were accompanied by the applicant's protest.

 No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US00/26884

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Group I, claim(s) 1-14, drawn to isolated nucleic acids and corresponding amino acid sequences for expansins (SEQ ID NOS: 1-6).

Group II, claim(s) 1-14, drawn to an isolated nucleic acid and corresponding amino acid sequence for a gibberellin-stimulated/ABA down-regulated cDNA (SEQ ID NOS: 7-8).

Group III, claim(s) 1-14, drawn to an isolated nucleic acid and corresponding amino acid sequence for an arabinosidase (SEQ ID NOS: 9-10).

Group IV, claim(s) 1-14, drawn to an isolated nucleic acid and corresponding amino acid sequence for a xyloglucan endotransglycosylase (SEQ ID NOS: 13-14).

Group V, claim(s) 1-14, drawn to isolated nucleic acids and corresponding amino acid sequences for cellulases (SEQ ID NOS: 17-20).

Group VI, claim(s) 1-14, drawn to an isolated nucleic acid and corresponding amino acid sequence for a polygalacturonase (SEQ ID NOS: 21-22).

Group VII, claim(s) 15, drawn to an isolated nucleic acid comprising a polygalacturonase promoter sequence (SEQ ID NO 23).

Group VIII, claim(s) 16-29, drawn to an isolated nucleic acid and corresponding amino acid sequence for a vacuolar H⁺ ATPase (SEQ ID NOS: 15-16).

Group IX, claim(s) 30-43, drawn to an isolated nucleic acid and corresponding amino acid sequence for an endo-β-mannanase (SEQ ID NOS: 11-12).

The inventions of Groups I-IX lack the same special technical feature in that the products differ structurally and functionally from one another, and the methods result in different products. The products of Group I are isolated nucleic acids encoding expansins, which is not a special technical feature of the products of Groups II-IX. The products of Group II are isolated nucleic acids corresponding to a gibberellin-stimulated/ABA down-regulated cDNA, which is not a special technical feature of the products of Groups I and III-IX. The products of Group III are isolated nucleic acids encoding an arabinosidase, which is not a special technical feature of the products of Groups I-II and IV-IX. The products of Group IV are isolated nucleic acids encoding a xyloglucan endotransglycosylase, which is not a special technical feature of the products of Groups I-III and V-IX. The products of Group V are isolated nucleic acids encoding cellulases, which is not a special technical feature of the products of Groups I-IV and VI-IX. The products of Group VI are isolated nucleic acids encoding a polygalacturonase, which is not a special technical feature of the products of Groups I-V and VII-IX. The product of Group VII is an isolated nucleic acid comprising a polygalacturonase promoter sequence, which is not a special technical feature of the products of Groups I-VI and VIII-IX. The products of Group VIII are isolated nucleic acids encoding a vacuolar H⁺ ATPase, which is not a special technical feature of the products of Groups I-VII and IX. The products of Group IX are isolated nucleic acids encoding an endo-β-mannanase, which is not a special technical feature of the products of Groups I-VIII. In addition, the methods of Groups I-VI and VIII-IX result in the production of different kinds of transgenic plants comprising isolated nucleic acids encoding different gene products. Therefore, lack of unity between the stated groups is properly made.

THIS PAGE BLANK (USPTO)

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER: _____**

**IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image
problems checked, please do not report these problems to
the IFW Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)